

# Are there tragedies without villains?

**On recent precedent, the latest crop of defendants in the French contaminated-blood saga will not be given a fair trial; some are already the victims of double-jeopardy.**

THE search for criminals to be held responsible for the French contaminated-blood scandal continues (see page 548). Among those now arraigned are M. Laurent Fabius, once President François Mitterrand's prime minister, and M. François Gros, the distinguished cell biologist and government adviser who happened to be the chairman of a crucial committee when the decisions to delay blood treatment and the screening of transfusion-blood were made in 1985. Everybody agrees, of course, that the decisions were wrong. Lives, perhaps hundreds of them, would have been saved if heat treatment and screening had begun earlier. It may also be the case that the decision eventually made was coloured by chauvinism, the preference for French rather than US reagents and techniques. But there are extenuating circumstances galore, of which the chief is the uncertainty at the time about the causation of AIDS and the reliability of the reagents on the market. Yet the trials that may yet be mounted, on charges of poisoning people, are unlikely to be fair.

That, at least, is the sad precedent set in the recent past. It will not be forgotten that four people have already been tried and convicted over the contaminated-blood affair. One of them, Michel Garretta, the former head of the French transfusion service, has spent two of four years in jail, while Jean-Pierre Allain, once head of research at the service, has only just been released from his term. Both are now to be tried again, but on charges of poisoning rather than of misleading consumers (of blood-clotting factor) by wrongly claiming that their products were safe. But each has already been tried twice; when Allain appealed against the original verdict, the French judicial system responded with a fresh trial, at which the judges could have decided on longer spells in prison.

Both convictions were unjust. Garretta, the head of the transfusion service, was by common consent the administrative head of a newly privatized public service with an incentive not to import supplies. By his own admission (in his defence), he had not kept up with research on AIDS. Allain, by the admission of the prosecution at his trial, had done much to warn Garretta of the dangers of contaminated blood. By the implication of his conviction, his fault was that of not rushing to *Le Canard Enchaîné* with his suspicions. But both were victims of the then-current climate of opinion and servants of the government whose officials are next in line for trial.

It is an offence against the civility for which France is distinguished that there should now be a fresh set of trials. Nobody should be surprised if observers now conclude that

the French judicial system will go on finding defendants for AIDS-related trials until the understandable anger of the relatives of those infected with HIV is stilled. Now that the new trials have been ordained, the best hope is that the court will address the question that really matters — not whether the decision taken in 1985 was right or wrong, but whether it was taken in good faith. If that is not the central question at the trial, France will have only itself to blame if cynicism about its judicial system is further increased. It is no part of the function of the courts to find a villain for every tragedy.

Meanwhile, the indictment of Fabius, two of his quondam ministers and Gros will be regarded with dismay by all those who now offer the French government advice on the conduct of its business, or who otherwise take responsibility for it as ministers. As modern technology goes, even in 1985 the supply of blood products must have seemed a relatively risk-free business. From civil aviation to nuclear power to biotechnology, there are other areas potentially more prone to danger, real or imagined. Nobody asks that advisers should be free to be irresponsible, but it will serve nobody's interest if they now habitually give ultra-safe advice for fear of otherwise being dragged through the courts when, in good faith, they give the wrong advice and are retrospectively identified as villains. □

## Plague in India?

**India's recent plague may have been something else, but the need remains to know more of the animal reservoir.**

SUGGESTIONS from India (see page 547) that the recent outbreak of what seemed pneumonic plague in Gujarat may not have been plague at all may be a great relief for India's tourist industry, but that does not necessarily mean they are incorrect. But it is too soon to be sure. The outbreak of disease in the city of Surat has certainly fallen short of gloomy expectations, but that could be a simple consequence of the speed with which people took avoiding measures. Alternatively, the outbreak may have been caused by an attenuated strain of the plague bacterium, perhaps one adapted to strictly respiratory transmission. Or the people of Gujarat may have a natural immunity. The plan of the World Health Organization (WHO) to mount an investigation in India is a good one. Let us hope that it reports as quickly as intended.

Meanwhile, it is not too soon to ask what the more distant



future holds. Plague is only one of several ancient infections lingering on the margins of human well-being. Tuberculosis and hepatitis-B are more prevalent and more serious problems of public health, as are malaria and trypanosome disease. But plague evokes atavistic fears even in comfortable Europe, where it has not been heard of since the seventeenth century. Curiously, there are even pockets of plague among rodents in the American southwest from which people occasionally die. Eradication is notoriously difficult when there is a reservoir of infection among a wild animal population, but that is no reason for not finding out. Instead of killing rodents as in the past few weeks, the people of India might be better employed on trapping them so that the prevalence of the plague bacterium in the wild might be estimated more accurately. □

## Magellan's success

**The spacecraft Magellan, due to end its sojourn around Venus on 11 October, is a success for NASA.**

THE Magellan mission to Venus, ended by the deliberate extinction of the spacecraft on 11 October, has had plenty to say both about Venus and the proper conduct of space science. At the most prosaic level, the \$900-million probe has successfully fulfilled its mission by sending radar signals through the planet's dense atmosphere, returning to Earth the most accurate geological map of any planet. More generally, the four-year mission is vivid testimony to people's resourcefulness. Cut back under the Reagan administration, the Magellan mission had to make do with a far simpler single-instrument probe than the National Aeronautics and Space Administration (NASA) had designed, yet it has proved to be of more use to more disciplines of science than anybody expected. The policy-makers, of course, may not be pleased with the observation of one project engineer that Magellan has demonstrated "the extraordinary variety of science that can be extracted from a single-string fiddle".

But Magellan was not exactly cheap. Its survival in the harsh conditions of an oscillation each 90-minute orbit between extremes of hot and cold has proved the worth of proper systems engineering. Ground controllers, for example, have been able to simulate the real effect of every command sent to Magellan throughout the mission by the use of a dummy probe on Earth. The advocates of smaller, faster and cheaper space exploration do not always bother with such niceties. That is one reason why their great hope — Clementine — knocked itself out in May. (The infuriating silence of the missing Mars Observer since 21 August 1993 is something else again.)

Even on its planned demise this week, Magellan was expected to glean extra information. The plan was to lower the orbit gently, causing a tendency for the probe to spin, and then to measure the viscosity of the atmosphere of Venus by the reaction of the booster motors as they adjust to prevent rotation — a memorable null experiment if ever there was one. That should be a guide to atmospheric composition. It

is also an object lesson in getting out while the going's good: to have let the probe struggle on until it failed would yield little science and plenty of bad publicity.

But perhaps the most impressive feature of the Magellan project is not its success as an engineering project, or even its prodigious output of scientific data, but the way in which data have been disseminated. Magellan's maps are already available to most of us on the Internet. Failing that, they can be bought on compact disc. Thus the data from the mission are not the plaything of some conclave, but are available to everyone to use. School science will have a field day when teachers tumble to what there is. In space science as elsewhere, that is the shape of things to come.

And the bad news? Magellan burned up leaves NASA's mispriorities even more exposed than before. A diminishing number of financially pressed but scientifically interesting planetary missions stand beside the vastly expensive manned space programme which masquerades as science but whose actual purpose lies elsewhere. Sadly, there seems not the slightest chance that NASA will learn from Magellan where to look for affordable success, and that it will invest more energy (and money) in the planetary programme at the expense of what is fast coming to seem the most expensive public works programme ever: space-station Freedom and its infrastructure. □

## Natural conservation

**Raptors, and falcons in particular, may have forfeited their right to conservation.**

EVERYBODY knows that nature is red in tooth and claw, but it can go too far. There will be consternation among readers of the London *Times* at the report (10 October) of recent happenings on the Ile Aux Aigrettes, off Mauritius. The island is one of the centres at which the Jersey Wildlife Preservation Trust, supported by the naturalist and writer Gerald Durrell, is seeking to nurse back to viability endangered bird species, notably the Mauritius kestrel and the pink pigeon once also indigenous to Mauritius. But now the unthinkable has happened. With both species now numerous enough to be returned to the wild after a captive breeding programme, it appears that a kestrel has taken a fledgling pigeon from its nest. The pigeon is by now presumed eaten.

Savagery of this kind puts conservationists everywhere in a fix. What purpose can be served by rescuing pigeons from the endangered list if they are then to be eaten by once-endangered falcons? Sadly, it is hard to see how this practice, instinctive though it may be, can easily be ended. Even if adult kestrels were conditioned against the consumption of pigeons before being returned to the wild, their offspring (absent Lamarck) would probably be less discriminating. For the long run, the outlook is even more gloomy. The small size of the population of surviving pigeons before their conservation was taken in hand may itself be a monument to the past activities of the raptors. There are hard choices to be made on the Ile Aux Aigrettes. □



# US adopts economy and civilian use as new goals for military research

**Washington.** The US Department of Defense (DoD) has announced that, for the first time, the affordability and possible commercial use of new military technologies will become key objectives of its \$36 billion-a-year research and development programme, not merely technological supremacy.

Describing this shift in emphasis in its new *Defense Science and Technology Strategy*, the DoD says that, although "technological superiority remains essential", it will also pursue two new goals, namely affordable weaponry and enhanced economic security.

But the department also says that it plans to "downsize, outsource and restructure" its science and technology infrastructure, including its own laboratories. Specific proposals will be submitted early in 1995 to the department-wide Base Realignment and Closure Commission.

"In the past we have focused on developing new capability, in some cases with not too much regard to cost," says Anita Jones, director of research and engineering at the DoD. "We can no longer do this on a reduced budget. We need to use technology to reduce the cost of systems. That is a new theme and a new objective for the science

and technology program."

Introducing the strategy document last week, Jones added: "We want to make the defense technology investment nurture those technologies that are the base for both military products and commercial products."

But John Deutch, deputy defence secretary and formerly dean of the Massachusetts Institute of Technology, sought to reassure

Their formal adoption in a strategy document at this stage is probably intended to help defend the science and technology programme from congressional budget-cutters — as well as some elements of the armed forces themselves — who regard it as extravagant and over-protected.

Last year, for example, a Senate appropriations committee described advanced technologies as representing "the un-affordable luxury of pursuing technology for technology's sake". But the DoD's research and development programme has been cut only marginally, while — as the strategy document notes — spending on weapons systems is being slashed by 60 per cent over a ten-year period.

Most of the \$36 billion to be spent by DoD this year on what

it calls 'research, development, evaluation and test' (RDT&E), will go on the development and trials of military systems. An estimated \$4 billion is spent on basic research and exploratory development. The universities spend \$1.8 billion of that — a figure that will fall by 10 per cent in the coming year — with the rest spread across industry and the DoD's own laboratories.

The new strategy offers five principles to guide the armed forces and defence agencies in managing science and technology: getting technology to the 'warfighters' more quickly; cutting costs; strengthening US industry; promoting basic research; and assuring quality.

The document also encourages DoD scientists to "monitor and collaborate in international science efforts". Jan Walker, a department spokeswoman, said that such collaboration would not be organized centrally, but was "in the hands of individual engineers and scientists".

A basic research programme is needed to ensure that the DoD "has early cognizance of new scientific ideas", the document says, pledging to "ensure stable funding for the highest priority efforts". Nevertheless, the document also commits the department to reducing its support for its own science and technology activities.

"The DoD must reassess the conditions under which it maintains in-house capability," the document says. "Today it may be more effective to rely on industry or universities for those technologies that are developing outside DoD at a rapid pace." But, as recent congressional reductions in DoD grants for university research demonstrate, such a shift in strategy may prove difficult to implement.

Colin MacIlwain



**Stealth bomber: technological showpiece, but was it cost-effective?**

the armed forces that the new strategy will not reduce the quality of the DoD's research programme. "The strategy has one overall goal, to maintain the technological superiority of the United States," he said. (The strategy document itself says that technological superiority "remains essential", even if it is no longer "sufficient".)

The ideas of cost-effectiveness and dual-use technology are hardly new to the DoD.

## Germany beats UK to XMM contract

**Munich.** The European Space Agency (ESA)'s Industrial Policy Committee (IPC) has decided to appoint Germany's DASA-Dornier as prime contractor for its XMM mission, a high-resolution X-ray telescope originally due to be launched at the end of the decade, although this date may be put back because of disagreements over the contracting process (see *Nature* 368, 87; 1994).

The decision over one of the four key 'cornerstone' missions of ESA's science programme was made despite the fact that a bid from British Aerospace received a higher assessment overall than that from Dornier, performing slightly less well on technical criteria, but better on management criteria.

A bid from Italy's Alenia Spazio was considered less attractive than either the German or British bids. The Italian company had led an earlier consortium bid for the ECU280-million (US\$338 million) contract, but failed because ESA felt it was too expensive and lacked acceptable proposals for programme management (see *Nature* 368, 87; 1994).

Space industry officials say that, in choosing Dornier over British Aerospace, ESA appealed to a rule, established by space ministers at a meeting in Grenada in 1992, that the 'industrial return coefficient' — the ratio of the value of contracts received by a country to its annual contribution to ESA — for each member state should not fall below 0.96 at the end of 1996, and should aspire to the ideal value of 1.0, embodying the concept of *juste retour*.

The IPC says that its decision will help to rectify the current imbalance of the industrial return coefficients. The United Kingdom had a coefficient of 1.04, representing a surplus of around ECU60 million a year. Germany's is 0.98, representing a deficit of around ECU100 million.

But ESA's decision could also lead Britain to raise further questions about the functioning and activities of the space agency itself, particularly as it is concerned that overhead and management costs are biting into the agency's budget for space science (see page 546).

Alison Abbott



# UK warns of cost short-fall over future space missions

**London.** The future involvement of the United Kingdom in the science missions of the European Space Agency (ESA) is being threatened by the fact that the increased cost of its subscription to the agency is undermining its ability to participate in individual space missions. Britain's difficulties could precipitate a full-scale review of the future of the European agency itself.

The warning comes from the newly formed Particle Physics and Astronomy Research Council (PPARC), which is keen to reduce the costs of its basic subscription to ESA — and, in the short term, to persuade the Office of Science and Technology to help it to accommodate current fluctuations in this cost.

Without additional short-term support (which could be made available when the science budget is allocated in December), PPARC has warned that British scientists may be forced to withdraw from participation in payload projects for ESA's next two missions, the gamma-ray observatory satellite INTEGRAL and the cometary probe Rosetta, both due for launch in 2001.

Britain's subscription to ESA has been rising because, although both the overall budget of the agency and that of the PPARC are nominally being kept level in 'real' terms, the former is being adjusted for a higher rate of inflation than the latter. One result is that the UK's subscription (like that to CERN, the European Laboratory for Particle Physics) is consuming an increasing proportion of PPARC's budget. "It is like being elected a member of a golf club, but having nothing left to buy balls and a club after paying your subscription," says Ken Pounds, professor of physics at the University of Leicester and chief executive of the PPARC.

In last year's white paper on science funding, the government promised to explore ways in which responsibility for fluctuations in international subscriptions could be spread across the whole science budget. Pounds has already met Sir John Cadogan, the director-general of research councils, to discuss how this might be achieved.

At their next meeting, Pounds will take with him the conclusions of the PPARC council, agreed to last month after a review of its scientific programme, that the council "would be unable to participate significantly in the national development of instruments for future ESA space science missions without additional funds from Government".

According to Pounds, these conclusions apply in particular to the INTEGRAL and Rosetta missions. Britain's main involvement in the first of these is through the

caesium iodide imager, one of the two main instruments to be carried on the satellite. As one of the four principal members of ESA — the others are France, Germany and Italy — Britain is hoping to play "a very significant role in both of these missions", says Pounds. "If we are not able to be involved, then it would send shock-waves through the European space programme."

## IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

### The Integral Mission: can Britain's astronomers still participate?

Roger Bonnet, head of ESA's science programme, admits that the agency is "very concerned" at the conclusions of PPARC's council, and accepts that a British withdrawal from INTEGRAL and Rosetta would have wide implications. "It would be a tremendous setback," says Bonnet, adding that he is already exploring ways of avoiding such a move.

Bonnet points out that ESA came out well from a recent review of its cost-effectiveness. But he admits that there may be scope for improvement, adding that, if the United Kingdom feels there are new areas worth investigating for potential savings, he would certainly consider doing so.

According to Pounds, Britain is not the only ESA member to be facing difficulties caused by the increasing cost of the ESA subscription. The only difference, he says, is that government restrictions on the science budget have meant that Britain "hit the brick wall first".

"In the short term, it is essential that the government offers PPARC protection against the currency fluctuations in our subscription to ESA — and to CERN," says Pounds. "In the longer term, we must work, with our international partners, to reduce the cost of these major science programmes."

Several countries are believed to feel that the problems facing ESA do not lie in the administration of its science programmes, but in the heavy infrastructure built around its efforts to develop a European manned-spaceflight programme, whose main incentive has evaporated with the ending of the Cold War.

Maggie Verrall

# Darwin deal marks growth of link-ups for gene companies

**San Francisco.** In the latest link-up between a large pharmaceutical company and a small genome-based biotechnology company, Rhône-Poulenc Rorer of France has signed a \$16.5-million agreement with Darwin Molecular Corporation, based in Seattle, Washington, to work on gene therapies for cancer.

The agreement combines Darwin's work on genes with the French company's vectors and gene-delivery technology, with the aim of targeting tumour cells. Darwin was established two years ago to convert information from the human genome into treatments for cancer, AIDS and autoimmune diseases.

Rhône-Poulenc will invest \$5 million in Darwin in exchange for about 10 per cent of its stock. It will provide the remaining \$11.5 million to Darwin over four years in the form of a second equity investment, research funding and milestone payments.

The deal follows several major investments by pharmaceutical companies in small genome companies. Last year, Human Genome Sciences Inc. of Rockville, Maryland, attracted \$125 million from SmithKline Beecham, and Millenium garnered \$70 million from Hoffman-La Roche. In June, Incyte, based in Palo Alto, California, reached a \$25-million agreement with Pfizer.

Darwin is believed to have been trying to set up a corporate collaboration for at least a year, with some observers claiming that the company has shown a lack of focus. But David Galas, acting chief executive of Darwin following the recent departure of Mark Pearson, says that the deal with Rhône-Poulenc Rorer is the first stage of what promises to be a larger commitment.

According to Galas, Darwin is in the early stages of talks with several companies about agreements on DNA sequencing, gene discovery and molecular evolution, a technology that uses environmental pressures to 'select' artificially new molecules with particular strengths.

Darwin and Rhône-Poulenc Rorer plan to develop a gene-therapy approach employing Darwin's novel herpes thymidine kinase genes to trigger the activity of the antiviral ganciclovir against tumour cells.

Rhône-Poulenc has said it aims to persuade biotechnology companies to work together in order to speed the development of new drugs and treatments, operating in a network similar to that of the software company Microsoft, whose founder William H. Gates is a key investor and board member of Darwin. Darwin would eventually be part of a larger constellation of biotechnology and pharmaceutical companies working with Rhône-Poulenc (see *Nature* 369, 92; 1994).

Sally Lehrman

● See also page 551

# Indian plague poses enigma to investigators

**New Delhi.** Doubts are emerging about the precise nature of the apparent outbreak of pneumonic plague in the Indian town of Surat. The World Health Organization (WHO) has launched an investigation into what it describes as a "very baffling and mysterious" outbreak, while there is scepticism among some Indian microbiologists about whether the disease is in fact pneumonic plague.

After a visit to India last week by Hiroshi Nakajima, director general of WHO, the organization formally announced on 10 October that a team of experts from the United States, Russia and WHO would investigate the outbreaks of both bubonic and pneumonic plague. The team, which includes epidemiologists, clinicians, a veterinary scientist and experts in communicable diseases, is expected to report within the next few weeks.

After visiting Surat and the laboratories of the National Institute of Communicable Diseases (NICD) in Delhi, Nakajima confirmed that there was pneumonic plague in Surat. But he added: "How it started, why all cases are pneumonic, and how it spared family contacts are just some of its features that are baffling."

But the unusual features of the outbreak have led some Indian microbiologists to question whether it really is pneumonic plague. N. P. Gupta, for example, a former director of the National Institute of Virology (NIV) in Pune, points out that pneumonic plague is a disease of cold climates, whereas the temperature in Surat was 35 °C.

Furthermore, contrary to experience elsewhere in the world, the alleged outbreak of pneumonic plague was not preceded by an outbreak of bubonic plague. Indeed, Bombay escaped without a single case, despite the influx of 50,000 panic-stricken migrants from Surat.

Enver Akhmedov, an epidemiologist at the Russian embassy in Delhi, is one of the foreign experts who admit to being surprised at the small number of confirmed positive cases in Surat, pointing out that "pneumonic plague is the most infectious of all known diseases".

Despite these doubts, WHO is convinced that no other disease is involved. "It is 100 per cent plague," said Nakajima.

Nakajima said his team agreed fully with NICD's identification of *Yersinia pestis*, the plague bacillus, as the organism involved in the Surat outbreak. "However, with only 196 out of 1.8 million population being positive for plague, I am hesitant to call it an epidemic," he added.

The fact that WHO's judgement is based on data presented to it by the NICD, the national centre for plague surveillance and control, has also upset the sceptics, some of whom have criticized the centre's techniques for identifying the bacillus. But specific

identification for *Y. pestis*, using fluorescent antibody test, was made by NICD on the day Nakajima visited its laboratory.

N. K. Shah, an expert adviser to WHO, says that the inquiry team is keeping an open mind, and will investigate all plausible explanations for the outbreak. The team will also look at the role of rats in the outbreak of bubonic plague in Beed in the state of Maharashtra, and the later outbreak of pneumonic plague in Surat in the neighbouring state of Gujarat. "We are 100 per cent sure of the involvement of rats in Beed," said Guneeal Rodier, an epidemiologist with WHO. "But Surat is the big question mark."

The Indian government is confident that both epidemics are now under control, and claim that the number of suspected new cases is continuing to decline. But the outbreak cannot be declared officially over until 12 days after the last case of pneumonic plague.

Subash Arya, a former deputy director of

NICD, says that much of the confusion over the cause of the outbreak is the result of the concentration of all plague-related work at NICD. But he also predicts an end to the dispute, as scientists at the All India Institute of Medical Sciences in Delhi have developed a DNA probe that can detect *Y. pestis* not only in humans, but also in samples from rodents and fleas.

The Indian Council of Medical Research (ICMR), which has no plague expert among its staff, has called for a meeting of all microbiologists — both retired or serving — who have worked with *Y. pestis* or handled a plague case. It has also directed the NIV to make an independent analysis of samples from Beed and Surat. ICMR's laboratory in Ahmedabad, the National Institute of Occupational Health, has taken air samples from Surat to detect any bacilli in aerosols, but has not announced its findings yet.

**K.S. Jayaraman**

## Nobel honours pursuit of G proteins

**London.** This year's Nobel prize for physiology or medicine throws cell signalling into the limelight. The prize has been awarded jointly to Alfred G. Gilman and Martin Rodbell to honour their discovery of the role played by G proteins in cellular signal transduction, a decision greeted as "tremendous news" by many scientists.

In 1971, Earl Sutherland received the Nobel Prize for discovering cAMP and the enzyme that creates it — adenylyl cyclase. At the time, it was assumed that when a hormone bound to its receptor at the cell surface, the receptor signalled directly to enzymes such as adenylyl cyclase in the interior of the cell.

As a result, when Martin Rodbell, then at the National Institutes of Health (NIH) (and now at the National Institute of Environmental Health Sciences), proposed in the early 1970s that there were intermediates between hormone receptors and the cell's interior, most researchers ignored their possible existence.

But in a series of ingenious experiments, Rodbell and Lutz Birnbaumer, his first postdoctoral fellow, demonstrated that the step between the hormone receptor and adenylyl cyclase was dependent on GTP (hence the name G protein). "Martin wrote out the protocol and I did the pipetting" remembers Birnbaumer.

The intermediate step had previously been missed simply because most scientists had unwittingly added GTP to their

experiments when using 'dirty' preparations, according to Robert Lefkowitz, professor of medicine and biochemistry at Duke University, who was then at the NIH.

Stephen Nahorski of the University of

Leicester in Britain describes Rodbell as a bright and generous scientist, always willing to discuss unpublished data, and keen to help younger colleagues. But Rodbell's observations also inspired Alfred Gilman, a young pharmacologist with a keen eye for important issues in pharmacology, to hunt for a GTP-dependent protein.

The isolation of the first G protein by Gilman (and independently by Thomas Pfeuffer) confirmed Rodbell and Birnbaumer's suspicions. It also left Gilman, now professor of pharmacology at the University of Texas Southwestern Medical Center, leading the field, as he has since done for many years. With relentless commitment, he has purified, cloned, and even crystallized G proteins, revealing the molecular basis of the way in which they transduce signals.

Rodbell's continuous production of ideas, combined with his infectious enthusiasm, have complemented Gilman's prolific experimentation and total commitment to his subject. Several hundred receptors that signal by this route are now known, and altered G-protein signalling plays a pivotal role in conditions like whooping cough and cholera, as well as various genetic conditions (see, for example, *Nature* 371, 164; 1994). **Harriet Coles**



**Gilman: analysed role in cell signalling**

AP/Wide World



# Blood scandal raises spectre of Dreyfus case

**Paris.** Science is on a collision course with the French legal system over the handling of investigations into the contamination of blood with HIV in the mid-1980s. Last week, the affair took a new twist with the decision to add François Gros, the former director of the Institut Pasteur, to the list of those indicted.

But the affair is turning into one of the biggest legal controversies since the historic 'Dreyfus affair' a century ago, when Alfred Dreyfus, an innocent Jewish army officer was convicted of treason. As in the Dreyfus case, France is deeply divided, this time between those who believe that the contamination is the collective responsibility of the state and those who see it as a crime perpetrated by a small number of individuals.

The divisions have sharpened following last month's indictment on charges of 'collusion in poisoning' of Laurent Fabius, the prime minister at the time, Edmond Hervé, the secretary of state for health, and Georgina Dufoix, the minister for social affairs.

Contributing to the controversy is the feeling that the truth of the matter remains elusive, despite repeated investigations and criminal proceedings. Indeed, Simone Veil, the health minister, declared last week that "a general explanation of the entire transfusion system in 1985" is still needed.

The French affair involves three separate issues. Were haemophiliacs given HIV-infected clotting factors when heat-inactivated alternatives were available? Was routine screening postponed to await a French test when a US test was already available? Were blood donations from prisons stepped up despite warnings about collecting from high-risk populations?

Axel Kahn, of the Cochin Institute of Molecular Genetics in Paris, says that one point of general concern is that the judiciary is conceding to the demands of a "vindic-

tive" public. "Guilt is being presumed," he says. The formula being used is "how dare you say you are innocent, that you did your best, when 1,000 are dead", says Kahn. "It's an extraordinary perversion of justice."

Contempt for the law "is the old French disease", said one lawyer, while another suggests that "to exorcise the big fear of our time, we need guilty parties". Jean Bernard, president of the National Blood Transfusion Centre (CNTS) board until December 1984 and former president of the national bioethics advisory committee, has also said that the courts should have called foreign scientists as witnesses, "because the same things happened everywhere".

Indeed, the Conseil d'Etat has fixed the state's responsibility "as of November 1983" on the grounds that the threat of HIV was then "established by the scientific community". But many argue that consensus was reached only after Robert Gallo published his rediscovery of HIV in April 1984.

More contentious is the Conseil's assertion that the efficacy of heat-inactivation was accepted by 13 October 1984. It based this on a recommendation — incorrectly attributed to the 'World Haemophilia Organization' — that "the value of the protection [of heated products] against the AIDS virus should be taken as established".

In fact, the recommendation, made by the US National Haemophilia Foundation, said only that clinicians "using coagulation-factor concentrates should strongly consider changing to heat-treated products with the understanding that the protection against AIDS is yet to be proven". Worryingly, both errors match previous ones made by a French journalist, suggesting that the Conseil did not consult the original documents.

Many of the investigations into the affair also have potential conflicts of interest. A 1991 inquiry was carried out by the General

Inspectorate for Social Affairs (IGAS), for example. But in a report published in May 1985, IGAS had endorsed the commercial approach of Michel Garretta, then head of CNTS. Running CNTS, it said, requires not skills in "research and teaching", but in "organization, methods and the promotion of technologies". Moreover, the word 'AIDS' does not appear once in the report.

Michel Lucas, the author of the 1991 IGAS report on the blood affair, also later supervised an IGAS report on the collection of blood in prisons that exonerated the ministry of justice. But Lucas had been co-president of the Prisons Administration when it issued a decree in January 1984 increasing the collection of blood from prisons.

Prison blood accounted for almost a third of all contamination in 1985, which partly explains why 1,300 people were contaminated through whole-blood transfusions in France, compared with 70 in the United Kingdom, where routine screening was introduced two months later than in France.

But the report claims, for example, that the Prisons Administration was unaware of a 1983 circular warning against collecting blood from high-risk populations, although the warnings were widely covered in the media. Health officials also allege that the ministry of justice demanded the deletion of an explicit reference to prisons in a 1985 revision of this circular; the ministry is said to have denied the allegation.

The report claims that instructions from the Prisons Administration in 1984 calling for an "increase in the frequency of collections... and in the number of donor detainees", was not intended to mean stepping up donations.

The question of whether officials delayed introducing routine screening using a US test because they were waiting for a French test is also more complex than is commonly portrayed. For one thing, it took place against the background of the bitter US/French dispute over the origin of the AIDS test.

And although the French were clearly protectionist, this seems to have been widely known, if not widely endorsed. In March 1985, the daily medical newspaper *Le Quotidien du Medecin* ran a headline "The War of the Tests", and the *Quotidien de Paris* (21 July 1985) reported that "for obvious reasons of national interest" the government prefers the Pasteur test.

Moreover, for at least part of 1985, there was widespread dissent about the wisdom of introducing routine screening (see *Nature* 367, 673; 1994). And although the US test had been approved by the US Food and Drug Administration (FDA) in early 1985, FDA documents support French arguments that there remained genuine concerns about the quality of the test.

**Declan Butler**

## Charges set a 'terrifying precedent'

**Paris.** François Gros, the leading French biologist and scientific adviser to the then prime minister, Laurent Fabius, during the mid-1980's, was last week indicted on charges of 'collusion in poisoning', along with Claude Weisselberg, adviser to Edmond Hervé, a former secretary of state for health.

The indictment of Gros has been viewed with widespread concern in the scientific community, as it appears to place the responsibility of science advisers in a new perspective. Axel Kahn, for example, of the Cochin Institute of Molecular Genetics in Paris, says that for scientists who get involved in political

IMAGE  
UNAVAILABLE  
FOR  
COPYRIGHT  
REASONS

Gros: indicted on poisoning charges.

decision-making out of civic duty, the charges set a "terrifying precedent" because it judges not the act but the consequences. Kahn adds: "It's equivalent to charging an 'inattentive' mother with 'homicide' because her child ran onto a road causing a crash that left 150 children dead."

Kahn, who is president of the committee that approves genetically modified organisms, says this could mean that he would face criminal charges, for example, if he approved a modified rape plant that later intoxicated people. Charges of 'collusion in poisoning' are "crazy, completely crazy".

Declan Butler



# Japanese earthquake tests disaster warning networks

**Tokyo.** Last week's major earthquake off the northern tip of Japan exposed both the strengths and weaknesses of Japan's system for preventing disasters in the aftermath of such events. Tsunami warnings were broadcast on television with impressive speed, and some accurate predictions were made. But information on the earthquake itself, which measured 7.9 on the Richter scale, was slower to emerge, and some of the information was inaccurate and misleading.

In Tokyo, where the earthquake was felt quite strongly despite being nearly 1,000 km from the epicentre (see map), the ground had barely stopped rumbling when television stations broke into their broadcasts to issue tsunami warnings.

Roughly five minutes after the earthquake, which struck at 10:23 p.m. local time, Japan Broadcasting Corporation (NHK) transmitted a map of Japan with the Pacific coast of the northern island of Hokkaido flashing in red to indicate a 'tsunami warning', and the Pacific coast of the main island of Honshu flashing yellow as far south as Tokyo to indicate 'tsunami caution' (a lower level of warning).

These represent the first two of a three-level warning system issued by the Japan Meteorological Agency (JMA). The third level of warning, 'great tsunami', indicates tsunami of a height of more than 3 metres.

The visual displays were accompanied by warnings in Japanese and English that residents on the Pacific coast of Hokkaido should flee to higher ground because tsunami "up to 2 metres high" might strike.

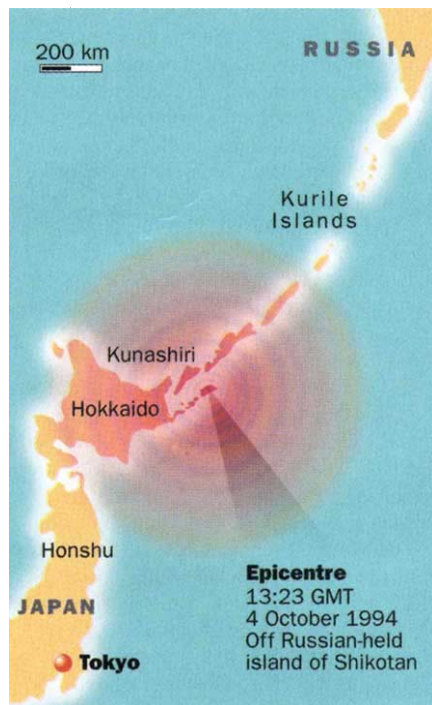
A few minutes later, the first predictions from JMA of times of arrival of tsunami were announced on NHK, with the earliest arrivals set at 11 p.m. (local time) for the northern tip of Hokkaido, 37 minutes after the earthquake. The first tsunami actually hit Nemuro on the northern tip of Hokkaido at 10:58 p.m. and was 1.73 metres high.

But after this initial success, things began to degenerate. At 11.34 p.m., more than an hour after the earthquake, the JMA upgraded its warning for the northern Pacific coast of Honshu to a tsunami warning rather than caution. JMA say they did this because of "the situation in Hokkaido". But this caused confusion among disaster prevention officials.

Under a tsunami caution, officials must go to the coast to observe the tsunami. Under a tsunami warning, where the tidal wave can be up to about 2 metres high, they must flee away from the coast. Some officials who followed the rules and went to the northern coast of Honshu under the tsunami caution, apparently did not get to hear the second upgraded warning one hour later. If a big

tsunami had come, says one critic, "they would have been zapped".

Broadcasts of information on the earthquake itself were even less impressive. The tsunami warnings began at 10.28 p.m., only five minutes after the earthquake. But they were not accompanied by any information about the earthquake itself, and television



viewers were left ignorant of where the earthquake occurred, or how large it was.

The first information on the size of the earthquake from JMA came several minutes later. It gave the magnitude of the earthquake in various cities on the Japanese scale of one (minor rumbling) to seven (a major disaster). The high readings, of up to six, were centred in northern Hokkaido.

Disaster prevention officials complain that such information is just as important as tsunami warnings for responding to an earthquake, and are calling on JMA to release it quicker. In response, JMA announced a few days later that, as from next March, it will announce the magnitude of earthquakes about two minutes after they occur, before the release of tsunami warnings.

Earlier this year, the agency added about 150 unmanned stations around the country for measuring the effects of earthquakes. But it has yet to organize this system to issue very fast measurements. From March next year, JMA says it will issue the highest readings for a given region, rather than waiting for all measurements from each region, within about 2 minutes of an earthquake.

David Swinbanks

## Former Italian minister accused over space funds

**Munich.** Sandro Fontana, Italy's former minister of research, has become the latest addition to the list of individuals indicted over the mismanagement of funds belonging to the Italian Space Agency (ASI).

Fontana, who was minister from July 1992 to August 1993, joins ASI's former president Luciano Guerriero and his board of directors, who were indicted in November last year partly on the charge that they had deliberately diverted funds away from basic research.

According to ASI's own rules, 15 per cent of its budget should be reserved for basic research. But three years ago a conflict between Guerriero and Remo Ruffini, chairman of ASI's science committee, which assesses research grant applications, over whether the 15 per cent should include contributions to the basic research activities of the European Space Agency (ESA), brought the agency to a virtual standstill (see *Nature* 356, 647; 1992).

To break the impasse, Fontana established a committee of five 'wise men' in February last year to investigate this issue and possible corruption in assigning contracts for ASI's controversial X-ray satellite SAX, which has still not been launched (see *Nature* 366, 101; 1994). Fontana gave the committee only one month to carry out its investigations, and then accepted its conclusion that there had been no wrongdoing at ASI, and that Guerriero's interpretation of the 15 per cent was correct.

The Court of Accounts has now ruled that Fontana should not have accepted the hastily completed report of the committee, which it considers to be flawed. It also says that Guerriero misled the committee by falsely interpreting a letter from Jean-Marie Luton, director-general of the ESA, as indicating that almost a fifth of Italy's 602 billion lire (US\$385 million) contribution to ESA in 1993 was spent on basic research.

Fontana, Guerriero and his board of directors have 30 days in which to reply to charges of mismanaging ASI's funds. If they are unable to provide the Court of Accounts with a satisfactory explanation, the court can decide either to impose a fine or to pass the case on to criminal courts to take legal action against those accused.

Meanwhile ASI itself is re-emerging from its period of political turmoil. Only one of those indicted is associated with ASI today (Enrico Cerrai is still a member of the board). A new board, under the chairmanship of Giorgio Fiocco, atmospheric physicist, was installed by Colombo in February this year, and Fiocco appointed a new scientific committee last month.

Alison Abbott

## Stanford hit by \$1m fines over hazardous materials

**San Francisco.** Stanford University has agreed to pay almost \$1 million in fines for mishandling hazardous waste materials, the largest fine ever levied on an educational institution in the United States. The university will pay \$460,000 in penalties to the state, \$235,000 in costs, and \$300,000 to three local environmental groups.

The university has now warned its researchers that, in future, they could be personally liable for fines arising from any infringement of state regulations for which they are found to be responsible.

Stanford officials have admitted liability for 40 per cent of 1,600 violations of which they had been accused, thus avoiding a lengthy court battle. The violations concerned toxic spills, mislabelling of containers, and inadequate storage both in university laboratories and at its Environmental Health and Safety (EHS) facility on campus.

"The problems we saw at Stanford were as widespread and serious as any problems we had seen anywhere in the state," says Allan Hirsch, a spokesman for the California Department of Toxic Substances Control. Hirsch claims that toxic materials are just as serious "in an academic institution as in a factory", and that the settlement puts other educational institutions on notice that they will be held to the same standards as industry.

According to Hirsch, discussions are already taking place between his department and the University of California at Berkeley, and California State University campuses at Hayward and Sonoma, over possible violations.

But Lawrence Gibbs, director of EHS, disagrees with Hirsch on the severity of the problems at Stanford. "Regulations are needed, but the requirements must be appropriate for the environment," he says, arguing that regulations based on conditions in industry are not directly applicable to the laboratory. For example, says Gibbs, industrial chemical operations generate large quantities of relatively few different types of hazardous waste. This is in contrast to research universities such as Stanford, which has approximately 2,300 sites that generate or store approximately 5,000 different kinds of chemicals, but usually in very small quantities.

Barbara Barres, a researcher in the university's department of neurobiology, says that this itself can create problems when it comes to labelling. Regulations supplied in a volume "a foot thick" stipulate that every chemical must be labelled with its full name, with no abbreviations allowed. In some instances this is impossible, says Barres, given the very small containers routinely used.

The present settlement involves violations between 1988 and 1992. But Stanford has had problems with its handling of hazardous materials since 1986, when a whistle-blower complained of design flaws in the university's new \$7 million EHS facility, including drains running in the wrong direction and an incinerator chimney that blew fumes back into the EHS offices.

In March 1992, the university told the state it had sorted out its problems. But an inspection one month later found hundreds of containers violating regulations; some were unlabelled or mislabelled, others were in a rusty condition, and many were improperly stored.

The state imposed 28 more citations on the university, saying that it had "intentionally or negligently falsified inspection records". More violations were uncovered in August and September 1993, and again in June this year.

A spokesman for the university claimed last week that the safety of both individuals and the environment "was not endangered". He said that three-quarters of the citations involved administrative processes, such as record-keeping, labelling, and reporting.

But some university researchers believe that the harshness of the fines which have been imposed on the institution may reflect the state's frustration with the way in which the university had attempted to deflect earlier criticism of its handling of hazardous materials.

Joel Shurkin

## Genentech agrees to drop promotion of growth hormone

**Washington.** In response to pressure from Congress, the biotechnology company Genentech has withdrawn its support for controversial 'height screening' programmes and other promotional activity for the synthetic human growth hormone Protropin, one of the first pharmaceutical products of genetic engineering.

The California-based company announced its move in a letter to Ron Wyden (Democrat, Oregon), chairman of the House of Representatives small business regulation subcommittee. Wyden has been investigating allegations that the screenings led to the \$20,000-a-year drug being prescribed for short children who are not hormone-deficient. Protropin, which earned \$216 million for Genentech last year, has been approved for the treatment of hormone-deficient children only.

The drug's distributor, Caremark, has also promised to cut all financial links with doctors who prescribe the drug. Caremark and Genentech executives are facing charges in Minneapolis — which they deny — of making illegal payments of \$1.1 million to a paediatrician who did so.

Wyden welcomed Genentech's move as "an effective step forward". But he added that it amounted to an admission that the promotional campaigns had gone too far, and promised to pursue tighter national standards to control the promotional activities of the pharmaceutical industry.

Colin Macilwain

## France boosts gene therapy centres

**Paris.** France is considering taking steps to simplify the regulatory procedures for gene therapy and to encourage the wider use of the technique in its public hospitals, according to Philippe Douste-Blazy, the country's junior health minister.

Gene-therapy trials were first approved by the government in 1990, when the national bioethics advisory committee concluded that somatic gene therapy posed no special ethical problems. But as protocols for gene therapy are handled under existing legislation, their approval falls under the jurisdiction of four different committees.

In practice, this bureaucratic labyrinth has been less difficult to negotiate than it might appear. A coordinated processing system, reinforced by the fact that the same individuals often sit on the different committees, has meant that more than half of the 16 protocols examined so far have been approved, all within less than six months.

With demand expected to increase, the government is now considering ways of streamlining the process. But rather than taking the long route of passing new legislation specifically aimed at gene therapy, it is considering creating an inter-ministerial commission that would take over the work of the existing committees.

The government also intends to attach small gene-therapy centres to hospitals across the country, in a bid to shift gene therapy techniques from the laboratory to the hospital ward.

Another longer-term proposition aimed at speeding up clinical trials and encouraging greater industrial investment in gene therapy is that the country should establish the equivalent of the US Orphan Drug Act. Under this act, which was passed in 1983, companies that develop drugs for rare diseases — defined as those that affect fewer than 200,000 people — are granted certain commercial and legal privileges.

Declan Butler



# Consortium plans 'public' map of genome

**London.** Last week's meeting in Washington to discuss strategies for compiling a transcript map of the human genome — a physical guide to the possible location of almost 100,000 human genes — agreed to co-ordinate efforts by an informal consortium of major research institutions from the United States, Britain and France to push the project forward (see *Nature* 371, 365; 1994).

Such a map is destined to become a basic research tool for the pharmaceutical industry in the next century. But members of the consortium, which is being organized under the umbrella of the Human Genome Organization (HUGO), are insisting that the results of their endeavours should be placed freely in the public domain.

At the Washington meeting, organized by Britain's Wellcome Trust, it was in particular agreed that all members must accept that the use of the transcript map by scientists from both the public and private sector should be "unencumbered" — that is free of any legal or commercial constraints.

One of the key participants in this consortium will be a research group carrying out a sequencing project at the Washington University in St Louis, Missouri, with support from the US pharmaceutical company Merck Sharp & Dohme.

Two weeks ago, Merck announced that, in a direct challenge to the conditions being required by Human Genome Sciences Inc. (HGS) on access to cDNA sequence data compiled by its Institute of Genomic Research (TIGR), headed by Craig Venter, the results of its own project would be made freely available to the scientific community with no conditions attached (see *Nature* 371, 463; 1994).

During last week's meeting, Merck's decision to back a 'public' sequencing exercise was enthusiastically endorsed by various scientists from academic institutions. Many expressed their concern over the terms of access to the TIGR sequence data being demanded by HGS and its chief sponsor, the pharmaceutical company SmithKline Beecham (see *Nature* 371, 365; 1994).

One participant, for example, said he had been told that no scientists would be allowed access to more than 100 partial sequences a year — equivalent to less than one per cent of the total database.

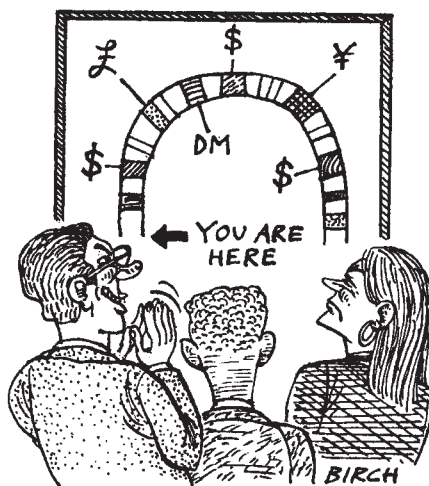
But Merck, which has endeavoured to stake out the moral high ground in the current dispute of public versus private 'ownership' of the human genome, also received the backing of several other pharmaceutical companies represented at the meeting.

Such companies claim to be worried that a small group of genome companies could corner the market in genome information — and thus potentially hold to ransom any other commercial enterprise seeking to use this information. "We feel that, although

patents may be granted on products containing [genetic] information, they should not be allowed on the information itself," says Barry Ross, research director of Glaxo. "My impression is that many other companies have the same feeling."

In response, representatives of HGS at the meeting are said to have used the occasion to defend their position vigorously, pointing out that several academic institutions, including the Howard Hughes Medical Institute, have already accepted the terms on which it is offering scientists access to TIGR's data.

William Haseltine, president of HGS, pointed out that TIGR has already sequenced more than 150,000 partial gene sequences, soon to be made accessible through its Human cDNA Databank. TIGR's activities had



already proved of immense potential value to the pharmaceutical and diagnostics industry, he said, adding that it was reasonable for those who had backed the project to be looking for a return on their investment.

Haseltine was supported in his defence of HGS's strategy by at least one other gene-sequencing company, Incyte of Palo Alto, California, which claims that it has already sequenced almost 30 per cent of the genome, and has entered into an agreement with at least one major pharmaceutical company (Pfizer Inc.) for access to its database.

Randy Scott, chief executive officer of Incyte, said that his company was in a position where, if funds were forthcoming, it could map "half of the genome" by the end of next year — providing that investors in such a project see the likelihood of a return on their money.

He also criticized the strategy of 'open access' advocated by Merck as potentially undermining the ability of information-based companies such as Incyte to help in the swift discovery of new genes through their sequencing efforts. "Merck's strategy could prevent anyone getting a patent on a complete sequence," Scott warns.

Despite the differences of opinion, it was agreed at the end of the meeting that members of the consortium, which are likely to include the Sanger Centre in Britain, the Whitehead Institute in Boston, and researchers at the University of Washington and Stanford University, would work towards pooling their efforts on a communal transcript map over the next 18 months.

A further meeting of the main participants will be held under the auspices of HUGO — probably in the United Kingdom — within the near future. This will discuss progress on efforts aimed, for example, at ensuring that participating institutions achieve common standards of quality.

Francis Collins, the head of the National Institutes of Health's (NIH's) Human Genome Programme, is reported to have expressed strong opposition during the Washington meeting to any private participation in building the transcript map which would (like the use of TIGR's cDNAs) involve constraints on the use of its data.

Collins has so far refused to comment further on NIH's plans. But it is expected that Harold Varmus, the director of NIH, will support his suggestion that the NIH should support work that is part of an international effort whose results are placed unencumbered in the public domain.

At the same time, a number of pharmaceutical companies — including the US company Pfizer and the UK-based Glaxo and Zeneca, all of which were represented at the Washington meeting — are also reported to be considering how, in support of Merck's strategy, they too could provide funds for a public-domain mapping project. "We would be prepared to support such an activity involving a partnership of industry and academic groups to map and sequence cDNAs," says Ross of Glaxo.

HGS has already indicated that it is holding discussions with other research groups on how a mapping project using the TIGR sequences might be carried out using private finance. But it will still be allowed to attend the next meeting of the 'public domain' consortium — even though it (in common with several other groups) is not at present prepared to accept the conditions required of full consortium members.

Michael Morgan of the Wellcome Trust says that part of the success of the heavily attended Washington meeting was the open dialogue it permitted between advocates of the conflicting approaches. But he remains adamant that the trust has little interest in backing efforts to map the genome with commercial strings on access to the resultant data. "The Wellcome Trust does not have a remit to protect industry," says Morgan, although adding that "at the same time, we have no wish to undermine it."

**David Dickson**



## Benefits of primary care

**SIR** — In an article "Primary care is not the answer" (*Nature* 370, 501; 1994), Barbara J. Culliton draws attention to some "less visible provisions" in President Bill Clinton's health-care plan, such as forcing 50 per cent of physicians into general practice. She believes that this "50 per cent solution" will fail, as modern medical practice is beyond the skills of the general practitioner. Medicine should be provided mainly by academic medical centres.

Primary care has always been the preferred way to practise medicine. Unfortunately, it has been abandoned by academic medicine. Diseases do not exist as such, there are only sick people, and yet our education systems produce mainly disease specialists. They may understand "basic biological phenomena", and know how to treat some diseases, yet the crucial question is not only how to treat but when to intervene. The latter is generally ignored because withholding treatment is regarded as unethical, and many patients therefore receive unnecessary treatment. Academic medicine turned *primum non nocere* into malpractice, and palliation into nescience. And yet most patients do not need more than that, seeking help from alternative medicine.

An important objective of Clinton's plan is to mobilize competent "academic generalists" and place them where they are needed most, in the community. To restore the lost dignity of the family physician who now faces new responsibilities, and to protect his patient from academic medicine. Medicine has become extremely complex and the family physician has to decide what is best for his patient in the same way as a lawyer does for his client.

The threat of academic medicine to the patient is real. The family physician has to protect him from unnecessary treatment, experimental remedies and clinical trials that study disease and ignore patient suffering.

Academic medicine is the main source of iatrogenesis, that is, drug sensitivity and antibiotic resistance (hospital strain). It is also the main cause of rising medical-care expenditure, as it indiscriminately applies technological innovations, when in reality the patient could benefit without them. As patient protector, the family physician will obviously also reduce medical expenses.

**Gershon Zajicek**

H. H. Humphrey Center for Experimental Medicine and Cancer Research, Hebrew University, Hadassah Medical School, PO Box 12272, Jerusalem 91120, Israel

**SIR** — In her diatribe against the role of primary care in Clinton's plans for health-care reform, Culliton demonstrates that C. Everett Koop does not have a monopoly on patent and dangerous nonsense. She misses the main point about general practice: namely that well-trained generalists act as vital gatekeepers for access to the specialists who are thereby allowed more time to practise their medicine on appropriate patients.

To take up her argument on "Koop's elbow", what if the pain is due to angina? Is the patient expected to know that in this instance a cardiologist would be better than a rheumatologist? Immediate access to specialist care in these circumstances is at best a waste of time and at worst dangerous, as life-saving treatment may be unnecessarily delayed. The United Kingdom's National Health care system may have many flaws, but, in the opinion of many doctors who work in it, a well organized network of highly trained general practitioners is not one of them. The United States would be well advised to emulate it.

**Cella Helllwell**

**Andrew Hall**

8 Park Avenue, Hexham, Northumberland NE46 3EN, UK

**SIR** — It was disappointing to see *Nature* dismiss efforts to reform the US health-care system (*Nature* 369, 508; 1994). As a molecular biologist whose livelihood depends partly on continued government spending for basic medical research, I share your concern that this source of revenue should not be lost.

However, as a parent unable to purchase medical insurance to meet the cost of treating a haemophiliac child, I am aware of severe shortcomings in our health-care system. A single dose of clotting factor now retails for \$1,300. Our health-care dollars support large numbers of workers in the insurance, legal, administrative, pharmaceutical, marketing and other professions.

In a widely quoted recent statement, the former US Surgeon General, C. E. Koop, points out that 72 per cent of Britain's physicians are primary care doctors, compared to 29 per cent in the United States. It seems unlikely that free market forces will ever reverse these trends. Unless elected officials play an active, positive role in restructuring the US health-care system, the United States will probably never achieve what other developed countries have, in being able to provide basic medical security for working citizens and their families.

**Marlana M. Goldrick**

513 Pigeon Forge Road, Pflugerville, Texas 78660, USA

## Record conditions for cycling

**SIR** — Miguel Indurain recently broke the world 1-hour cycling record, riding 53.040 km. The previous record was held by Graeme Obree, who rode an innovative bicycle of his own design and manufacture. Subsequently, the UCI (Union Cycliste Internationale) changed its rules to outlaw such innovations. The UCI does not, however, regulate the atmospheric pressure or composition used for the record attempt. That is why both the Merckx (1972) and Moser (1984) hour records were achieved at altitude in Mexico City, where the air is less dense than at other tracks. Indeed, both Obree and Indurain have plans to attack the record in Mexico City next year.

One of us recently proposed attempting the hour record in a depressurized indoor track filled with  $\sim 0.2$  atm of pure oxygen<sup>1</sup>. This approximates to the atmosphere used in modern spacecraft, and has the same volumetric oxygen content as air. At world record speeds, frictional losses from tyres and drive train become an insignificant fraction of total resistance<sup>2</sup>.

We can also assume that the drag coefficient will not change, a good assumption for bluff bodies at high Reynolds numbers<sup>3</sup>. For constant power output, the speed will then scale as density<sup>(-1/3)</sup>. For a 0.2 atm O<sub>2</sub> atmosphere, the speed increases by a factor of  $(0.2 \times 32/29)^{(-1/3)} = 1.655$ . The 32/29 factor accounts for the higher molecular weight of O<sub>2</sub> versus the average air molecule.

Almost any competitive cyclist could then break the current record (including either one of us) within the current UCI rules. Obree or Indurain could ride more than 87 km in one hour to take the title definitively.

The hour cycling trial is one of the most difficult of all sporting events, and those currently contesting it are arguably the fittest of all humans. Thus the proposed experiment would test the absolute limits of human physical performance under space flight conditions.

**Jonathan E. Snow**

CNRS/CRPG,

BP20,

54501 Vandoeuvre-les-Nancy, France

**Mark Drela**

Department of Aeronautics

and Astronautics,

Massachusetts Institute of Technology,

Cambridge, Massachusetts 02139, USA

1. Snow, J. & Shames, P. *UK Cycling Weekly*, No. 5332, p. 27 (1994).

2. Whitt, F. & Wilson, D. G. *Bicycling Science* (MIT Press, Cambridge, 1983).

3. Hoerner, S. *Fluid Dynamic Drag* (Hoerner Fluid Dynamics, 1967).

# Trade and the biodiversity convention

Daniel M. Putterman

**How can the international trade in genetic resources help global biodiversity conservation? The UN's Biodiversity Convention provides the means to resolve this question.**

THE practice of screening tropical natural products for novel pharmaceuticals has been much in the news lately. With the adoption worldwide of the UN Convention on Biological Diversity, the interests of the pharmaceutical and biotechnology industries seem to have been pitted against those of the developing world. The convention has spotlighted the practice of collecting 'genetic resources' from developing countries, and researchers in industrialized countries, whose careers depend on the free flow of these resources, are upset by uncertainty over future access to them. Nevertheless, and despite appearances to the contrary, the convention is not an attempt by conservationists to lock up the world's genetic resources behind a wall of preservationism. Quite the contrary, it is meant to promote world trade in these resources, should result in more research and development, and deserves the cooperation of the international research community.

How can an environmental treaty also be a trade document? Articles 6–14 outline a broad plan for designing national conservation priorities, with due attention given to biodiversity preservation, sustainable use, conservation research, public education and the use of environmental impact assessments. These are, of course, crucial goals, but none of them is really new, in the sense that conservationists have long been advocating global attention to these issues.

## More money

Funding for these activities will come from the Global Environment Facility, a pre-existing fund replenished recently with another 3-year tranche worth more than \$2 billion, administered by the World Bank for the purpose of funding the "incremental costs" of global environmental protection. Incremental costs are defined as the additional costs of environmental protection that provide global benefits above and beyond those expenditures which ensure only local or national benefit. Although the Biodiversity Convention calls for the identification of "new and additional" financial resources to help developing countries meet these incremental costs, the Global Environment Facility has been designated as the interim funding mechanism. At international meetings of the signatories to the convention, developing countries have asked for new money to implement these

goals. But will the industrialized world set up another biodiversity fund?

Much of the Biodiversity Convention is thus a strategy for spending the Global Environment Facility money. This is a landmark achievement, as for the first time around 170 countries have acknowledged the common goal of biodiversity conservation, and have agreed on how to proceed. But the Global Environment Facility, like all development assistance, suffers one main drawback. Developing countries can wrangle over how best to spend the money, carving a slice here for biodiversity, a slice there for the amelioration of global climate change, and yet another sliver for the protection of international waters, but at the end of the day they are still dependent on the largesse of rich countries. The funding pie is fixed.

The most innovative — and controversial — articles of the Biodiversity Convention deal with access to genetic resources, to technology in general, and to biotechnology. Article 16 in particular refers to patented technology, and to other classes of intellectual property. Why did those who drafted the convention risk the objection of powerful industrialized countries like the United States by referring to intellectual property?

The world is moving toward freer international trade, promoted by international agreements such as the General Agreement on Tariffs and Trade and the North American Free Trade Agreement. Most developing countries have made industrialization one goal for economic development, with a priority being the acquisition of foreign technology. What can these countries trade for technology from the industrialized world? In the past, they have offered access to large consumer markets and cheaper labour pools, or have traded agricultural goods and natural resources.

Article 15 of the Biodiversity Convention recognizes the "sovereign rights" of nations over their genetic resources. It suggests that countries will soon add genetic resources to their list of tradeable goods, and hints at a new class of resource rights, the right to ownership of natural genes and chemicals. Genetic resources are not the "heritage of mankind", they are the property of the citizens of the countries where the overwhelming majority of species reside. This recognition of sovereign rights is intended to encourage world trade in genetic resources, the new

"comparative advantage" of the gene-rich yet technology-poor developing world.

Because of the vague wording in the convention about trade and technology, there was at first considerable opposition in the United States from the drug industry, which feared compulsory licensing of intellectual property. This issue was defused late last year when the White House issued an "interpretive statement", and the industry is to be commended for moderating its views on the convention, even cautiously embracing it.

## Tradeable goods

Promoting trade in genetic resources is the true genius of the Biodiversity Convention, for it does not limit funding for conservation to a few fixed development-assistance programmes. The goal of the convention is not merely to offer guidelines for slicing up development assistance funds for conservation, it is to enlarge the pie. Along the way, natural-products screening, or "biodiversity prospecting", will create new economic incentives for biodiversity conservation.

In the global effort to turn genetic resources into tradeable goods, the actions of individuals, whether corporate or academic, remains problematic. Much of the needs of industry for natural products are still supplied by a network of private brokers. Much of this 'underground' trade thrives on the ignorance of scientists and policymakers from developing countries of the value of their genetic resources. Theoretical arguments that the marginal value of genetic resources is too low to provide an incentive for biodiversity conservation are misguided and ignore the fact that each natural product has a trade value of tens if not hundreds of dollars. Adding value to these samples by chemical or biochemical means can increase their value at least tenfold.

The cooperation of researchers who collect biological samples in developing countries is essential. Scientists who sell samples to multinational companies without prior informed consent undercut the position of developing countries in international trade, and this is perceived as a serious threat by governments desperate to revive moribund economies. Increasingly, individuals are being investigated or even arrested for 'suspicious activity', so great is the mistrust of foreign researchers in some places.

There is a danger of throwing the baby



out with the bathwater. If the perceived threat to their sovereignty over genetic resources is great enough, governments can — and in some cases already have — prohibited natural-products collection. What is most needed now is education to assist policymakers in developing countries to design policies that guard against the unauthorized use of genetic resources by collectors while allowing governments to feel comfortable about maintaining access to their biodiversity.

### Quick fix

A quick fix which developing countries can resort to immediately is to insist on the use of material transfer agreements defining the scope of research on biological samples shipped abroad. Developing countries could incorporate these agreements into the collection permits which they routinely issue to researchers, retaining certain rights to commercialization and sale of products derived from the samples collected. A more sophisticated approach, first proposed by Janzen *et al.*<sup>1</sup>, calls for the use of detailed “research contracts” to define collaborative research projects with host-country scientists involving genetic resources. Such contracts would include provisions to share the benefits of natural products research equitably between all parties.

Some governments are already examining these policy questions. Spurred by scientific interest in a vine endemic to southwestern Cameroon, *Ancistrocladus korupensis*, from which scientists at the US National Cancer Institute have isolated an alkaloid that inhibits the growth of human immunodeficiency virus *in vitro*<sup>2</sup>, the Cameroonian government has decided to take a closer look at medicinal plant research. The prime minister's office recently formed a committee to look into strategies for conservation and commercial exploitation of *A. korupensis*, though what the country really needs is a strategy — of its own devising — for future research with all genetic resources. Cameroon's initiative presages that of other cash-strapped tropical countries.

Right now, the only example of this process is in Costa Rica. By law, all Costa Rican species, with the exception of flora found on private land, are part of “national patrimony”. The government has granted a non-exclusive license to develop genetic resources to INBio, the well-known national biodiversity institute which made news in 1991 when it signed a biodiversity prospecting contract with Merck & Company. The government taxes INBio's private-sector deals to fund conservation in the national parks, neatly solving the problem of who benefits from trade in genetic resources.

Of course, Costa Rica is well-known for the honesty of its government, and people have rightly questioned the applicability

of this example to more corrupt administrations. It is easy to imagine that in many countries a few people would benefit from genetic-resource trade at the expense of the majority. As with any development project, there is no easy answer to the problem of corruption and greed, or the inequitable application of human rights. Ultimately, the degree to which trade in genetic resources benefits biodiversity conservation will depend on the willingness of governments, corporations and private citizens to play fairly. It is to be hoped that private conservation organizations will work hard to encourage this.

Policy issues aside, just how large is the market for genetic resources? A look at multinational industries which already rely to some extent on tropical biodiversity for the development of new products reveals that the combined world market for pharmaceuticals, agrochemicals and seeds exceeds \$250 billion annually<sup>3,4</sup>. Capturing even as little as 1% of this market for developing countries would exceed the entire budget of the Global Environmental Facility.

Questions of sovereignty over genetic resources need not pit the interests of industry against those of developing countries. The drug industry's fears that article 15 will choke off the flow of genetic resources for new drug research are overstated. Given that the process of natural-products screening is a lottery, with the odds of discovering marketable products less than one in ten thousand (true at least for the pharmaceutical industry), it can be argued that gene-rich nations — assured that they will benefit from this trade — will have an incentive to supply greater quantities of natural products to industry.

Developing countries are already encouraging the formation of strategic alliances with pharmaceutical or biotechnology companies in the United States, Europe and Japan. In turn, some companies have found that these research partners offer a convenient, cost-effective solution to the problem of acquiring high-quality natural products and associated value-adding research services. Encouraged by the success of INBio, which has raised more than \$3.5 million in the past 3 years through joint ventures in the natural-products arena, and encouraged too by Australia's efforts to develop Amrad Corporation, a state-run but quasi-independent biotechnology company specializing in marine natural-products screening, scientists and conservationists in developing countries are nurturing home-grown natural-products ventures across the globe.

One example is the Bioresources Development and Conservation Programme, founded in Nigeria by Maurice Iwu, a medicinal chemist with substantial research experience in the United States, which has now spread to five Central

African countries. The programme seeks to commercialize genetic resources for the purpose of economic development at the community level, and to encourage resource conservation. Because traditional knowledge is part of Africa's “comparative advantage” on the world scene, the programme includes traditional healers, farmers and other rural Africans in the search for marketable new products.

Iwu decries the practice of what he calls “safari science”, pointing out that the process of drug discovery from tropical natural products would be more cost-effective if pursued largely in source countries. The Bioresources Development and Conservation Programme has a cooperative research and development agreement with Shaman Pharmaceuticals, and more contracts are pending. Many other such ventures are in the works. Recognizing the scarcity of start-up capital for such nontraditional business ventures, Thomas Eisner of Cornell University has proposed the creation of a “biotic exploration fund” to provide start-up grant money for 10 INBio-like “chemical prospecting” laboratories around the globe. Even the International Finance Corporation of the World Bank has floated a proposal for a “biodiversity venture capital fund” to provide start-up money for biodiversity prospecting, among other endeavours.

### Winner's charter

The use of genetic resources as a lever in the world marketplace holds a powerful appeal for developing countries. Yet many scientists still seem to underestimate this appeal. Merely attempting to introduce the subject of sovereignty over genetic resources caused rancorous debate at last year's meeting of the American Society of Pharmacognosy, and the issue was shelved entirely for this year's meeting. The society plans to hold an “interim” meeting in Costa Rica later this month to discuss the sovereignty issue.

Amid all the heat and smoke surrounding the national sovereignty clause of the Biodiversity Convention, scientists and industry alike are in danger of losing sight of its ‘win-win’ approach to research and development. Developing countries are becoming aware of the true value of their genetic resources. Natural-products researchers would do well to recognize this, and to prosper by it. □

Daniel Putterman is at the US Agency for International Development, Rm 509, SA-18, Washington, DC 20523-1812, USA.

The views expressed here are those of the author and not those of USAID.

1. Janzen, D. H. *et al.* in *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development* (eds Reid, W. V. *et al.*) 131–157 (World Resources Institute, Washington, DC, 1993).
2. Manfredi, K. P. *et al.* *J. med. Chem.* **34**, 3402 (1991).
3. *Scripps*, 32 (January 1994).
4. World Bank Tech. Pap. No. 133 (Washington, DC, 1991).

# Academics in business mix poorly

**Academic institutions have been unwarrantably slow in coming to grips with the their members' involvement with commercial companies, hazarding their main goals in the process.**

THE age of our innocence was a long time ago, but that does not absolve the academic research community from the charge of cynicism in the pursuit of the financial benefits of research. The question of what is proper has come to a head again with the fuss in Washington last week about access by academics to the sequences of human gene tags accumulated by two US companies, The Institute for Genome Research (TIGR; proprietor J. Craig Venter) and Human Genome Sciences (HGS) Inc., with financial and scientific support from SmithKline Beecham. But it is a recurring dilemma.

In 1981, for example, over lunch in Cambridge, Massachusetts, Mark Ptashne, the Harvard professor, outlined one version of it. At the time, Ptashne's colleague Walter Gilbert was on leave from the university to function as president of Biogen Inc., then based in Geneva. Ptashne's ambition was to set up a company that would do similar things, but perhaps better. (Genetics Institute, as the company became, has indeed been a success.) Ptashne's difficulty was how to acknowledge the university's implicit but unquantifiable contribution to anything that might come of his plan. His solution was to offer 10 per cent of the equity to Harvard. Harvard declined the offer.

More than 30 years earlier, at the University of Manchester, England, the same issue arose in a different context, and was instinctively fudged. The late F. C. Williams, professor of electrical engineering, had interested Ferranti Ltd in building an electronic computer, an advanced machine for 1947. Long before the idea that software might enjoy copyright protection (and before the notion of an operating system had surfaced), the late Alan Turing dragooned those in the university who could into writing elementary library programmes.

In all the excitement, nobody seemed to worry about his or her intellectual property. For one thing, it was fun. For another, it was generally understood that if the enterprise succeeded, the university would be blessed with a decent computer science department (which happened anyway). And Ferranti would no doubt have become even more generous with loans and gifts of electronic and microwave equipment. It is a pity that so little came of the enterprise.

That was the age of innocence. Academics knew that their first responsibility was to students, their second to research and that it was also in the public interest that they should do what they could for industry, preferably local industry, which could then

be tempted to support the university as a whole. The understandings were implicit, but almost bankable. Now, there are contracts and memoranda of understanding about the ownership of intellectual property.

That is inevitable, but raises difficulties for academic institutions and also for governments, now almost united in their exhortation to academics that industrially relevant bright ideas are the best. The trouble is that, while the climate has changed, the goals of academic institutions have not. Are not students the first claim on an institution's responsibility? And then research whose temper will improve the quality of teaching and even be part of it (where graduate students are concerned)? As external subventions shrink, of course, there are many legitimate worries that academically purposeful research will be unduly cut back.

What, in those circumstances, should guide an academic institution's discharge of the general responsibility of the research community towards industry? The obvious first principle must be that nothing should interfere with the education of students, nor with research planned to that end. After all, successful industries usually say that they look to universities first and foremost for able and well-educated young people. (It is the lame ducks who look for patent licences instead.) Yet the conditions that must be satisfied by academics' industrial activities help to ensure that results are hardly ever discussed at universities.

Instead, there are discussions on questions such as the amount of time an academic researcher may legitimately spend working as a consultant for an industrial company, for a government or even for an international agency. Academic equality then ensures that the agreed rule will apply to all. Even the question of whether a consultancy will impair a person's capacity to give impartial advice to all and sundry is not taken up. Indeed, the nature of consultancy agreements is not habitually disclosed, certainly not to the close colleagues who may, nevertheless, be called upon to take a class that an over-burdened consultant-academic has to abandon at short notice.

But that is the small change of academic business. The road to riches demands more daring. To found a commercial company is a now commonplace route. But academic institutions are curiously and even culpably innocent of the implications. It is not simply that chairmen, chief executives and board members of companies have overriding fiduciary obligations to their shareholders

that can be potentially overwhelming of their academic responsibilities, but that the affairs of a new company are an incorrigible thief of the imagination of all those caught up in them. At least for a time, the academic entrepreneur will be a less dedicated teacher.

Yet few academic institutions have thought it worthwhile to have formal regulations on this subject. That is in stark contrast to what happens in commercial life, where an employee of one company can be a director of another only with the formal consent of his or her employer. The objective is to avoid intolerable conflicts of interest or intolerable external demands on time and energy. Are academic researchers somehow exempt from these conflicts? Certainly not in the opinion of those companies now said to maintain databases of the commercial connections of prominent academics in biotechnology and molecular genetics.

This is where government exhortations often ring false. Newly founded companies are small businesses, and small businesses are the engines of economic growth, even though most of them fail. So the climate has changed, and academic entrepreneurship is blessed. Nobody seems to have calculated the damage that may indirectly be done to the quality of teaching on which business, or at least successful business, also depends.

These complaints do not apply directly to research contracts let by industrial companies to academic departments, at least when the arrangements for the publication of the outcome do not prejudice the impartiality of the academics and students who are involved. There, the question is whether the project is one that will provide students with the challenge and the experience they need in preparation for a research degree. It may, for example, be a matter of argument whether a person who works out a quality-control routine in the manufacture of silicon chips can fairly earn a PhD in physics. Given the general eagerness of academic institutions for external income, even these questions are too little discussed.

The mess in molecular genetics has arisen because there are so many opportunities, because they seem to lie just within reach and because the rewards of occasional success appear to be so great. How can academic institutions hope to hold the line against such temptations? There is, of course, a market solution: buying out the would-be entrepreneurs. Sadly, few institutions are well-placed enough to follow it, but that does not mean that they can let the problem find its own solution.

**John Maddox**



# Chernobyl, eight years on

Dillwyn Williams

THE explosion at the Chernobyl nuclear reactor was an unprecedented disaster that has put the health of millions of people at risk. This is one version of events. The opposite view has it that Chernobyl was indeed a serious accident, but that although a lot of radiation was released it was quickly diluted and will have little discernible effect on the health of the exposed population. The truth, of course, is likely to lie somewhere between these two extremes.

Eight years after the accident, how do things stand? Meetings held in the past few months, and papers published, support the view that there has been a considerable rise in the incidence of thyroid carcinoma in children in the area around Chernobyl, but that there has not been a great increase in the reported incidence of other tumours or of genetic defects. Radiation is a mutagen and, like other mutagens, may exert its therapeutic effect on tumours in part through deficient repair of tumour-cell DNA. As a mutagen it has been associated with an increased risk of a wide variety of malignancies ranging from brain tumours to skin tumours and sarcomas. Two of the best documented radiation-induced malignancies are leukaemia and thyroid carcinoma.

To consider leukaemia first. Exposure to radiation is a known cause of the disease, and children are especially sensitive. No increase has been found in the incidence of childhood leukaemia in the areas around Chernobyl<sup>1</sup> so it is reassuring but hardly surprising that no increase has been found in the Scandinavian countries<sup>2,3</sup> or in broader European studies<sup>4</sup>. That does not prove that no increase will occur (and the studies so far have not had the power to detect a small increase) but it makes it unlikely that there will be any major effect. A report suggesting that an isolated increase in births with trisomy 21 in Berlin was related to exposure to fallout<sup>5</sup> differs from larger European studies<sup>6</sup>, and we do not know in how many other European cities an enterprising geneticist or paediatrician has scanned the records of births with severe genetic defects and failed to find an association with fallout from Chernobyl. If no effect is shown in an area of high exposure it is hardly likely that any effect will be shown in areas of low exposure, with a few exceptions where the radiation dose is so high that it sterilizes cells rather than causing them to become neoplastic.

In contrast to the studies of leukaemia or genetic defects, analyses of the incidence of thyroid cancer in children have been conducted mainly in the areas around Chernobyl. The first announce-

ments of an increase in incidence<sup>7,8</sup> have been followed by further documentation of the pathology<sup>9,10</sup>. The numbers reported have continued to increase and the tumours have been reported in northern Ukraine as well as in southern Belarus, and more recently in that part of the Russian Federation close to Chernobyl. These are a coherent group of results, in that the highest incidence and shortest latent period appears to be in the area exposed to the highest level of fallout. But they are not based on unbiased epidemiological studies, and the contribution of screening for thyroid tumours to the reported incidence is not known.

The strength of the association between the incidence of childhood thyroid cancer and exposure to fallout from Chernobyl was underlined this summer at two meetings\* in Nagasaki. Continuing increases in the numbers of childhood thyroid cancer diagnosed were reported, bringing the total in Belarus, since Chernobyl, to 251 and in the Ukraine to 276. Most of the Belarussian cases have occurred over the past four years, in the Gomel region, closest to Chernobyl, where the rate of childhood thyroid cancer reported has reached over 100 per million children per year. In most national tumour registries the childhood thyroid cancer rate is between 0.5 and 3, so the scale of the reported increase is dramatic.

The one problem preventing unqualified acceptance of the reported figures is that of ascertainment. Small papillary carcinomas, often only a millimetre or two across, are common incidental autopsy findings in adults. However, the pathology of the tumours reported from the countries involved is not that of microscopic lesions but of aggressive cancers, which are almost all papillary in type and have often directly invaded tissues around the thyroid. At one of the meetings, the result of a straw poll of an 11-strong panel of international experts was unanimity that the diagnosis of the tumours had been adequately substantiated and that exposure to the Chernobyl accident was definitely or probably the cause of the increase. Ten of the 11 considered that increased surveillance was of medium or low importance in contributing to the size of the increase.

The most likely cause of this increase in childhood thyroid cancer, with no apparent major increase in leukaemia or other tumours, is the radioiodine content of the fallout. This contrasts with the safety of <sup>131</sup>I treatment of adults with Graves dis-

ease. The limited growth potential of the normal follicular cell<sup>11</sup> may well be relevant to this discrepancy. The presence in fallout close to Chernobyl of short-lived isotopes of iodine as well as <sup>131</sup>I may also be important; the amount of <sup>131</sup>I released was almost a third of that of <sup>131</sup>I. Evidence of the link is likely to be indirect; the dose to the thyroid was measured after the event in only a small proportion of those exposed, and reconstruction is complex.

Molecular epidemiology may help. The sequence of events in thyroid carcinogenesis is partially known, and for papillary carcinoma *ret* or *trk* translocations, *met* overexpression and *ras* point mutations are among the commoner findings. Radiation is more likely to cause deletions or translocations than most chemical mutagens, and a limited study<sup>12</sup> has looked for *ret* translocations in childhood thyroid cancer in the Chernobyl area. The small number of cases successfully studied and lack of information on the frequency of *ret* translocation in childhood papillary carcinoma makes it difficult to draw conclusions. A different pattern of mutations in radiation-induced as opposed to non-radiation-induced experimental thyroid tumours has been reported<sup>13</sup>. Clearly, it is of considerable interest to compare the pattern of oncogene involvement in childhood thyroid carcinomas exposed to fallout from Chernobyl with a non-exposed control series.

Finally, study of the consequences of Chernobyl should be put into perspective. The increase in childhood thyroid cancer is relatively large, but it is not yet a serious problem in overall public health terms. Nonetheless, evidence from those exposed to external radiation or to radiation during the nuclear tests on the Marshall Islands shows that cases of thyroid cancer can occur for decades after exposure. Moreover there is a chance that the combination of human error and ageing reactors will cause a similar disaster. Analysis of the consequences of Chernobyl is essential to inform the response to any comparable event in the future. □

Dillwyn Williams is in the Department of Histopathology, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK.

- Ivanov, E. P., Tolochko, G., Lazarev, V. S. & Shuvaeva, L. *Nature* **365**, 702 (1993).
- Auvinen, A. *et al. Br. Med. J.* **309**, 151–154 (1994).
- Hjalmar, U. *et al. Br. Med. J.* **309**, 154–157 (1994).
- Parkin, D. M. *et al. Eur. J. Cancer* **29A**, 87–95 (1993).
- Sperling, K. *et al. Br. Med. J.* **309**, 158–161 (1994).
- Little, J. *Paediatr. perinat. Epidemiol.* **7**, 121–151 (1993).
- Kazakov, V. S., Demidchik, E. P. & Astakhova, L. N. *Nature* **359**, 21 (1992).
- Baverstock, K., Egloff, B., Pinchera, A., Ruchti, C. & Williams, D. *Nature* **359**, 21–22 (1992).
- Furmanchuk, A. W. *et al. Histopathology* **21**, 401–408 (1992).
- Nikiforov, Y. & Gnepp, D. R. *Cancer* **74**, 748–766 (1994).
- Wynford-Thomas, D., Stringer, B. M. J., Harach, H. R. & Williams, D. *Experientia* **39**, 421–423 (1983).
- Ito, T. *et al. Lancet* **344**, 259 (1994).
- Lemoine, N. R., Mayall, E. S., Williams, E. D., Thurston, V. & Wynford-Thomas, D. *Oncogene* **3**, 541–544 (1988).

\*Chernobyl Update and Chernobyl in the Future, Nagasaki, Japan, 3–4 June 1994. Proceedings to be published as Excerpta Medica International Congress Series 1074.

# Secrets of a double-doughnut

F. Ulrich Hartl

THE high-resolution structure of the molecular chaperone protein GroEL from *Escherichia coli*, presented by Braig *et al.*<sup>1</sup> on page 578 of this issue, will cause a stir in the many laboratories interested in the process of cellular protein folding. As an immediate application, the structural information will now allow the knowledge-based mutational analysis of GroEL. Indeed, the accompanying study by Fenton *et al.*<sup>2</sup> on page 614 demonstrates how such an approach can provide detailed mechanistic information.

GroEL is a member of the chaperonins, a family of molecular chaperones with ATPase activity that mediate protein folding in bacteria, chloroplasts and mitochondria<sup>3-5</sup>. More distantly related family members are present in the eukaryotic cytosol and in archaeobacteria. GroEL was initially characterized more than two decades ago as a host factor required for bacteriophage assembly<sup>5</sup>. Its general role in the folding of newly synthesized polypeptides has been established only recently, forcing a revision of the long-held view that protein folding and assembly in the cell occur spontaneously<sup>3-6</sup>.

Although the information specifying a protein's three-dimensional structure is encoded in its amino-acid sequence, it is now clear that a variety of pre-existing proteins are needed to effectuate this information in the intact cell. As demonstrated *in vivo*<sup>7</sup> and *in vitro*<sup>8</sup>, the chaperonins increase the yield of folded protein by suppressing the incorrect intra- and intermolecular interactions of unfolded or partially folded polypeptide chains that can lead to aggregation<sup>3</sup>. These off-pathway folding reactions are favoured by the highly concentrated intracellular environment, and the chaperonins use a fascinating mechanism to help other proteins to avoid them: unfolded polypeptide binds into the central cavity of the GroEL cylinder, and is thereby shielded from the cytosol. Subsequently, several rounds of ATP-dependent release and rebinding result in correct folding. GroEL cooperates in this process with the single-ring co-factor GroES.

The general features of the GroEL–GroES reaction cycle, as well as a number of mechanistic details, have been established by biochemical studies<sup>3-5,9</sup>. Now, thanks to the elegant work by Braig *et al.*<sup>1</sup>, the structure of GroEL has been defined to 2.8 Å. The authors obtained chaperonin crystals of high quality by massively overproducing GroEL in *E. coli*, enabling them to prepare very pure protein. However, they were also aided by a fortuitous twist of fate. Upon amplification of the *groEL* gene by the polymerase

chain reaction, two accidental point mutations — substitution of arginine by glycine, and alanine by valine, at positions 13 and 126 respectively — were introduced into the 547-residue sequence, resulting in a fully functional protein whose crystals diffracted very well. The structure<sup>1</sup> confirms the general architecture of the chaperonin established by electron microscopy<sup>10-12</sup>. It shows the unliganded

GroEL as a porous cylinder of 14 subunits, of relative molecular mass ( $M_r$ ) 58,000, composed of two heptameric toroids stacked back-to-back with dyad symmetry (Fig. 1).

The subunits of GroEL are folded into three distinct domains<sup>1</sup>, the largest being the well-ordered equatorial domain, containing residues 6–133 and 409–523 which form seven tightly packed  $\alpha$ -helices. The equatorial domain provides most of the intratoroidal side-to-side contacts and all the contacts between the two rings. It also contains the ATP-binding site (a structure of ATP-bound GroEL is in preparation<sup>1</sup>). A small intermediate domain (residues 134–190 and 377–408) connects the equatorial with the apical domain through short antiparallel polypeptide segments, which probably serve as hinges allowing allosteric domain movement. The intermediate domain also forms a diagonal projection that contacts the neighbouring apical domain on the right and with its base flanks the ATP-binding site. The apical domain (residues 191–376) forms the opening of the central cavity. A number of segments from this domain facing the central channel and the top surface are structurally ill-defined<sup>1</sup>. The mutational analysis<sup>2</sup> shows that these flexible regions are involved in the binding of polypeptide substrate and GroES.

Perhaps the most intriguing feature of GroEL is its large central cavity, which is known to accommodate the unfolded polypeptide<sup>10-13</sup>. Although it is unobstructed in the crystal structure, the cavity is probably subdivided by the five amino-terminal and the 23 carboxy-terminal residues of the subunits which are not crystallographically resolved and which protrude into the channel at the level of the equatorial domains<sup>1</sup>. It is unclear whether these — presumably flexible — segments prevent the transfer of folding polypeptide between rings<sup>9,11</sup>. For the most part, the channel has a uniform diameter of ~45 Å, but at the level of the intermediate domains it expands outwards to a maximum of ~90 Å (Fig. 1b). As well as the top and bottom openings of the cylinder, there are seven elliptical windows (36 × 13 Å at their widest extension) that connect the outside to the bulges in the central channel<sup>1</sup>. These windows may serve as entry and exit routes for nucleotides; they may also extend the available peptide-binding surface and reduce the steric restraint to allosteric domain movement.

In the accompanying study, Fenton *et al.*<sup>2</sup> used the structure of GroEL as the framework for an extensive mutational analysis of the chaperonin system. They describe a number of mutant forms of GroEL that affect the binding of one or more of its three ligands: unfolded polypeptide, GroES and ATP (Fig. 1). Interestingly, the mutations causing a

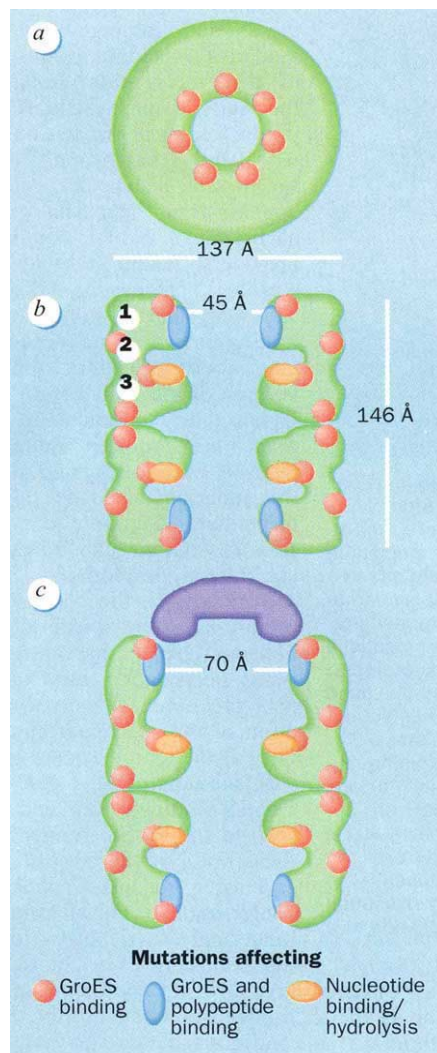


FIG. 1 Representation of the GroEL double-ring cylinder<sup>1</sup> displaying the functionally relevant regions identified by mutagenesis<sup>2</sup>. *a*, Top view into the central channel. *b*, Vertical section through the cylinder (modified according to Fig. 1 of ref. 1). 1, 2 and 3, the apical, intermediate and equatorial domains, respectively. *c*, Vertical section through the GroEL–GroES complex. GroES binding causes outwards movement of the apical domains in the adjacent GroEL toroid and less pronounced changes in the opposite toroid<sup>12</sup>. Locations and dimensions of regions of functional interest are approximate (see Fig. 3 of ref. 2 for details).



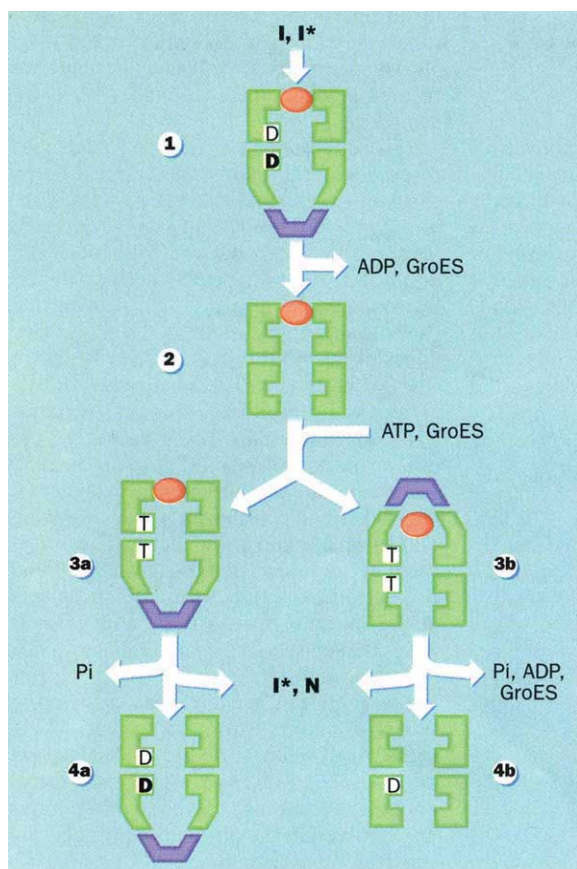


FIG. 2 Model for the ATP-dependent interaction between GroEL, GroES and folding polypeptide. **D**, the high-affinity ADP state in the seven subunits of GroEL which are bound to GroES; **D**, the lower ADP affinity of the subunits in the opposite toroid which may hydrolyse ATP (ref. 9); **T**, the subunits in a GroEL toroid in the ATP-bound state; **I**, polypeptide substrate as compact folding intermediate; **I\***, folding intermediate, partway advanced towards the native state (for polypeptides with separate domains or subdomains); **N**, native protein. (1) Polypeptide binding facilitates dissociation of tightly bound ADP and GroES. (2) Polypeptide bound stably in nucleotide-free toroid. (3) ATP and GroES rebound. GroES associates either with free GroEL ring (3a) or with polypeptide-containing ring (3b). ATP-hydrolysis leads to polypeptide release (4a) and incompletely folded polypeptide rebinds in (1). In 3b, polypeptide is transiently enclosed in the central cavity and is free to fold. ATP-hydrolysis in the GroES-bound toroid then generates the tight ADP-state (not shown). ADP dissociates upon ATP-hydrolysis in the opposite ring, causing dissociation of GroES and allowing polypeptide release. Polypeptide may rebound in (2).

defect in polypeptide binding mapped exclusively to the less-well-resolved segment of the apical domain facing the central channel. This region is probably composed of two  $\alpha$ -helices and an underlying loop segment<sup>1</sup>. Remarkably, changing any one of eight highly conserved, hydrophobic residues (two tyrosines, one phenylalanine, three leucines and two valines) in this region to glutamate or serine was sufficient to abolish polypeptide binding<sup>2</sup>.

Significantly, this confirms that the polypeptide binding site is in the chaperonin cavity<sup>10–13</sup> and sheds new light on the possible properties of the interaction. GroEL is thought to bind unfolded

polypeptide as a compact folding intermediate or 'molten globule'<sup>8,14</sup>. These flexible conformational states transiently expose hydrophobic surfaces, predisposing them to aggregation. It seems plausible that the residues identified by Fenton *et al.*<sup>2</sup> provide an adjustable, hydrophobic binding surface around the inside of the chaperonin cylinder that would be well suited to associate with complementary surfaces on the bound folding intermediates. A sliding motion of the polypeptide along this surface may facilitate partial folding in association with the chaperonin<sup>8,15</sup>. From the structure it appears that the cavity in each ring of GroEL could provide sufficient space for a protein of at least  $M_r$  35K, assuming that a compact folding intermediate occupies about 1.7-times the volume of the native state. Given that the apical domains are probably quite flexible, both locally and at the junction with the intermediate domain<sup>1</sup>, considerably larger substrates could be accommodated.

Strikingly, the same mutations in the apical domain of GroEL that abolished polypeptide binding also prevented the association of the co-chaperonin GroES, a heptameric ring with subunits of  $M_r$  ~10K. GroES binding to one end-surface of GroEL reduces the affinity of the opposite end for a second GroES, conferring structural and functional asymmetry to the double-ring<sup>9–12</sup>. Simultaneous binding of GroES to both ends of the GroEL cylinder has been observed under certain conditions, but the functional significance of this remains unclear (see ref. 16). GroES is required for the release of tightly bound polypeptide from its attachment sites at GroEL by coordinating the ATPase activity of the chaperonin<sup>3</sup>. Upon ATP-binding or hydrolysis, conformational changes in the equatorial domains may be conducted by the intermediate segments to the apical domains, causing them to bury their polypeptide-binding regions.

The demonstration that the binding sites for polypeptide and GroES overlap<sup>2</sup> now raises the interesting possibility of a

release mechanism by direct physical displacement, perhaps involving the previously described flexible loop segment of GroES<sup>17</sup>. This is also consistent with the demonstration by cryoelectron microscopy<sup>12</sup> that GroES binding induces a massive outwards movement of the apical domains of GroEL, creating an enclosed, dome-shaped space with a maximum height and width of 70 Å (Fig. 1c). GroES could initially make contact with the outer surface of the GroEL cylinder, triggering further domain movement and thus transiently displacing the polypeptide substrate into the cavity for folding. Indeed, Fenton *et al.*<sup>2</sup> identified other sites at the surface of the apical domain that were essential for GroES binding but did not affect the interaction with unfolded polypeptide.

Several other mutations affecting GroES binding mapped to the intermediate domain of GroEL, to the ATP-binding site and even to the interface between the two toroids<sup>2</sup> (Fig. 1). This may reflect the highly dynamic nature of the GroEL–GroES interaction and the intensive allosteric communication between the GroEL subunits during the reaction cycle<sup>9</sup> (Fig. 2). Nucleotide binding and hydrolysis, coordinated by GroES, switch the GroEL toroids between high- and low-affinity states for unfolded polypeptide. In the absence of substrate, GroES cycles slowly between bound and free states, alternating between the two end-surfaces of GroEL<sup>9,16</sup>. The co-chaperonin interacts preferentially with ATP-bound GroEL and, following ATP-hydrolysis, stabilizes the subunits of the interacting toroid in a tight ADP-state<sup>9</sup>. A subsequent round of ATP-hydrolysis in the opposite ring then causes dissociation of ADP and GroES (not shown in Fig. 2). Polypeptide binding to the toroid that is not occupied by GroES accelerates both ATP-hydrolysis<sup>8</sup> and dissociation of ADP and GroES<sup>9</sup>. Rebinding of GroES may then occur either to the polypeptide-bound or the

1. Braig, K. *et al.* *Nature* **371**, 578–586 (1994).
2. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. *Nature* **371**, 614–619 (1994).
3. Hendrick, J. P. & Hartl, F. U. *Rev. Biochem.* **62**, 349–384 (1993).
4. Ellis, R. J. *Curr. Biol.* **4**, 633–635 (1994).
5. Georgopoulos, C. *Trends biochem. Sci.* **17**, 295–299 (1992).
6. Ellis, R. J. *Nature* **328**, 378–379 (1987).
7. Horwich, A. L. *et al.* *Cell* **74**, 909–917 (1993).
8. Martin, J. *et al.* *Nature* **352**, 36–42 (1994).
9. Martin, J., Mayhew, M., Langer, T. & Hartl, F. U. *Nature* **366**, 228–233 (1993).
10. Langer, T., Pfeiffer, G., Martin, J., Baumeister, W. & Hartl, F. U. *EMBO J.* **11**, 4757–4765 (1992).
11. Saibil, H. R. *et al.* *Curr. Biol.* **3**, 265–273 (1993).
12. Chen, S. *et al.* *Nature* **371**, 261–264 (1994).
13. Braig, K., Simon, M., Fumy, F., Hainfeld, J. F. & Horwich, A. L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3978–3982 (1993).
14. Hayer-Hartl, M. K., Ewbank, J. J., Creighton, T. E. & Hartl, F. U. *EMBO J.* **13**, 3192–3202 (1994).
15. Gray, T. E. & Fersht, A. R. *J. molec. Biol.* **232**, 1197–1207 (1993).
16. Todd, M. J., Viitanen, P. V. & Lorimer, G. H. *Science* **265**, 659–666 (1994).
17. Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. & Gierasch, L. M. *Nature* **364**, 255–258 (1993).

empty GroEL toroid, resulting in ATP-dependent protein release for folding (Fig. 2). Both mechanisms of protein dissociation may coexist. The extent to which folding occurs before the polypeptide is discharged into the bulk solution remains to be determined. Several rounds of binding and release of incompletely folded polypeptide have been observed both in the absence<sup>8</sup> and presence of GroES (see ref. 2).

With an accurate structure of GroEL in hand, progress towards understanding the complex interactions underlying the path-

way of chaperonin-mediated protein folding will be greatly accelerated. Further structural analysis will be required to elucidate the conformational consequences of GroES and nucleotide binding to GroEL, and the exact topology of the unfolded polypeptide within the central cavity of the chaperonin cylinder. □

*F. Ulrich Hartl is in the Howard Hughes Medical Institute and the Cellular Biochemistry and Biophysics Program, Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA.*

## QUASARS

# Lessons from a lucky star

Dennis Downes

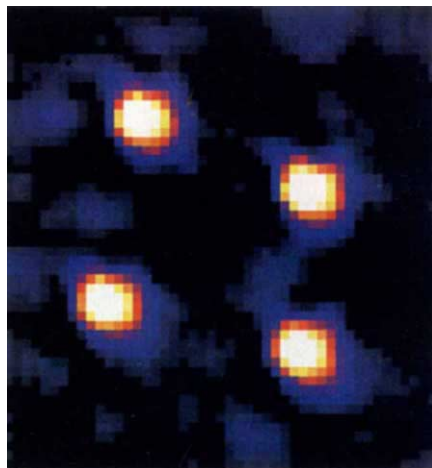
THE remarkable quasar H1413+117 bears the nickname of the 'Cloverleaf' because six years ago, optical astronomers<sup>1</sup> discovered that its light was coming from four spots on the sky — the result of gravitational deflection by a massive, unidentified galaxy or pair of galaxies along our line of sight. This celestial four-leafed clover certainly brought luck to Barvainis and co-workers who, on page 586 of this issue<sup>2</sup>, report their detection of carbon monoxide gas in the host galaxy of this quasar, at a redshift of 2.5 — the most distant molecular gas yet detected.

In the Cloverleaf, the carbon monoxide (3–2) line, normally at a rest frequency of 345 GHz, is redshifted by the expansion of the Universe to a frequency of 97 GHz. This redshifted line has been found with both of the telescopes of the French-German-Spanish Institute for Millimetre Radio Astronomy: the IRAM Interferometer on Plateau de Bure, France, and the IRAM 30-m telescope on Pico Veleta, Spain. This double detection, plus corroborating CO and neutral carbon lines seen at the 30-m telescope, make the identification rock solid. This is a welcome success, coming as it does after recent hints of CO lines in other quasars which unfortunately have not been confirmed by subsequent observations.

The interest stems from hopes of finding radio lines of molecular gas in ever more distant objects which resemble the long-sought *primaeval* galaxies in that much of their mass ( $10^{10}$  to  $10^{11}$  solar masses) is still in molecular gas, not yet condensed into stars. After CO gas was observed<sup>3–6</sup> in the galaxy IRAS F10214+4724<sup>7</sup> at a redshift of 2.3, searches were started at millimetre- and submillimetre-wavelength telescopes all over the world for CO in objects at high redshift. The Cloverleaf detection represents the first fruits of that search.

The CO line gives an accurate redshift for the host galaxy of the quasar, and

shows once again that the catalogued redshifts of quasars may not be those of their host galaxies. Rather, the standard high-excitation quasar emission lines often have a Doppler shift, due to outflow in a jet or wind, superposed on the



Optical (V-band) view of H1413+117, the 'Cloverleaf' quasar, from a 900-s exposure taken in March 1992 with the Planetary Camera on the Hubble Space Telescope. An unidentified galaxy along our line of sight deflects the light from the quasar into four bright spots, separated by 0.7 to 1.2 arcsec. North is towards the lower left of the picture, and east is towards the lower right corner. Courtesy of E. Falco, Smithsonian Astrophysical Observatory.

cosmological redshift. In the Cloverleaf quasar, the high-excitation emission lines are blueshifted by  $1,000 \text{ km s}^{-1}$  relative to the newly found CO line. The broad lines of highly ionized atoms seen in absorption against the quasar's continuum radiation are even more blueshifted, by a wind or jet streaming out of the quasar at  $12,000 \text{ km s}^{-1}$  (refs 8, 9). In contrast to this hot, fast-moving gas, the CO line comes from a large mass of neutral, cool, molecular gas extended over more than several hundred

parsecs. It therefore tells us the true redshift of the host galaxy.

When the Cloverleaf quasar was first discovered<sup>10</sup>, it was recognized as belonging to the class of broad-absorption-line (BAL) quasars. The new detection of CO gas shows that some BAL quasars are embedded in host galaxies with extraordinary amounts of molecular gas, about five times as much as the  $\text{H}_2$  mass of any other known galaxy and 100 times more  $\text{H}_2$  gas than in our own Milky Way. From the CO and dust detections, this gas seems to have carbon and oxygen at about the same abundance ratios to hydrogen that are observed in the Milky Way and in nearby galaxies, presumably because of an early phase of vigorous star formation.

This is the first detection of CO from a gravitationally lensed object, suggesting that amplification by a gravitational deflector may make molecular line detections a little easier. One wonders if the exceptional luminosity of the distant infrared galaxy F10214+4724 might also be partly due to gravitational lensing. However, the intensity boost from gravitational lensing is probably not the key to this new detection. The Cloverleaf spectrum has narrow absorption lines at redshifts of 1.4 and 1.6, from otherwise invisible galaxies about two-thirds of the way along our line of sight to the quasar at redshift 2.5. Models<sup>11</sup> for the Cloverleaf, based on lensing by one or both of these galaxies, say that the lens magnifies by a factor of seven the small, 1-light-day region emitting the quasar continuum and the 1-light-year zone emitting highly ionized, broad lines in the optical. As the CO line from the cool molecular gas has quite a low surface brightness, it must extend over at least a thousand light years to be detectable at all. The CO is thus a fuzzier source for the gravitational lens, and its signal is unlikely to have been amplified as much as the very compact optical emission.

More significant than amplification by an intervening gravitational lens is the radiation flux at submillimetre wavelengths. This was the clue that led the observers to search for CO. The Cloverleaf is the only radio-quiet quasar with a high submillimetre continuum flux<sup>12</sup>, which is the thermal radiation produced by an enormous mass ( $10^8$  to  $10^9$  solar masses) of dust, mixed with the even greater mass of molecular gas. This is the recipe for future line searches: concentrate on objects where you find submillimetre or far-infrared emission from dust.

How does the Cloverleaf manage to be one of the brightest of the BAL quasars if it is surrounded by so much dust? Barvainis *et al.* suggest an extension of the unified model for Seyfert 1 and Seyfert 2 galaxies: the Cloverleaf is a 'quasar 1' — that is, it has the usual high-ionization, broad lines. The big molecular ring



around the Cloverleaf is seen face-on, so the ultraviolet light from the quasar can get out to us without being stopped by the heavy curtains of dust. The authors think that the galaxy IRAS F10214+4724, on the other hand, is a 'quasar 2', with narrow emission lines, high polarization due to dust scattering, and a molecular ring seen more edge-on, blocking the light from a hidden quasar.

What next? The early Universe beckons. After 30 years of studying quasar and radio galaxies via their energetic matter moving relativistically or in highly ionized states, astronomers can now use the methods of millimetre and submillimetre astronomy to study the massive, cold (20–80 K), non-ionized material in high-redshift quasar host galaxies. The line and continuum results in H1413+117 and IRAS F10214+4724 were obtained with telescopes having collecting areas of about 1,000 m<sup>2</sup>. We can expect the discoveries to come thick and fast once the next-generation array pushes the collecting

area up to 10,000–20,000 m<sup>2</sup>. Thanks to the molecular Cloverleaf discovery, the design plans for such an array are indeed being born under a lucky star. □

Dennis Downes is at the Institut de Radio Astronomie Millimétrique, 300 Rue de la Piscine, Domaine Universitaire, F-38406 Saint Martin d'Hères, Grenoble, France.

1. Magain, P. *et al.* *Nature* **334**, 325–327 (1988).
2. Barvainis, R., Tacconi, L., Antonucci, R., Alloin, D. & Coleman, P. *Nature* **371**, 586–588 (1994).
3. Brown, R. L. & Vanden Bout, P. A. *Astrophys. J.* **397**, L19–L22 (1992).
4. Kawabe, R., Sakamoto, K., Ishizuki, S. & Ishiguro, M. *Astrophys. J.* **397**, L27–L30 (1992).
5. Solomon, P. M., Downes, D. & Radford, S. J. E. *Astrophys. J.* **398**, L29–L32 (1992).
6. Tsuboi, M. & Nakai, N. *Publ. astr. Soc. Japan* **44**, L241–L245 (1992).
7. Rowan-Robinson, R. *et al.* *Nature* **351**, 719–721 (1991).
8. Drew, J. E. & Boksenberg, A. *Mon. Not. R. astr. Soc.* **211**, 813–831 (1984).
9. Turnshek, D. A., Foltz, C. B., Grillmair, C. J. & Weymann, R. J. *Astrophys. J.* **325**, 651–670 (1988).
10. Hazard, C., Morton, D. C., Terlevich, R. & McMahon, R. *Astrophys. J.* **282**, 33–52 (1984).
11. Kayser, R. *et al.* *Astrophys. J.* **364**, 15–22 (1990).
12. Barvainis, R., Antonucci, R. & Coleman, P. *Astrophys. J.* **399**, L19–L22 (1992).

## DEVELOPMENTAL BIOLOGY

# The limb bud – part two

Malcolm Maden

THE first part of this story came in the form of a News and Views article<sup>1</sup> 15 years ago, in which Jonathan Slack described contemporary knowledge of pattern formation in the chick limb bud. He could discuss only two areas of the limb bud — the progress zone and the zone of polarizing activity (ZPA) — and not a single gene was mentioned. Contrast this with the string of genes and their varied expression domains which now have to be taken into account (see figure), and which show both how far we have come and how far we have to go in understanding how a limb is formed. The complex chain of cause and effect is being unravelled, however, as papers in *Cell*<sup>2</sup> and on page 609 of this issue<sup>3</sup> attest.

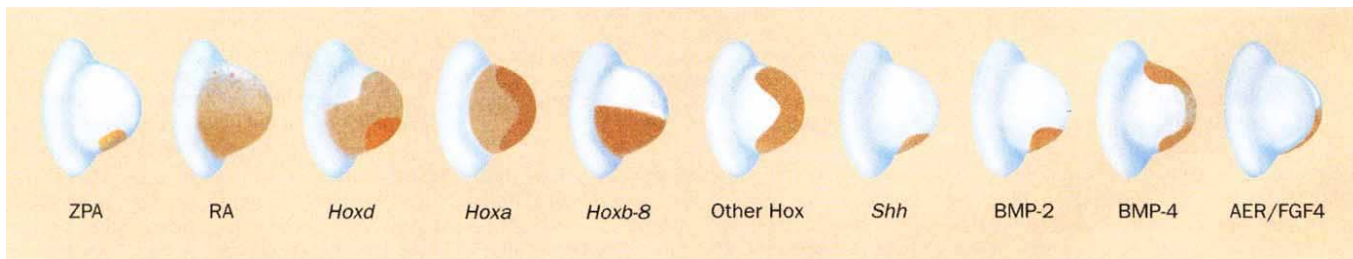
The problem remains the same as 15 (or 150) years ago: how is a highly complex limb made from a few thousand homogeneous mesenchymal cells covered by a thin ectodermal sheet? Up to now, answers have been obtained in two ways. First, by seeing how an extra limb can be induced; second, by examining the expression patterns of developmental control genes such as the homeobox (Hox) genes. In these ways it has been found that the players in the normal limb are the ZPA; retinoic acid (RA); products of various Hox genes; the product of sonic hedgehog gene (*Shh*); bone morphogenetic proteins (BMPs); the apical ectodermal ridge (AER); and fibroblast growth factor (FGF). There

will surely be others that have yet to be identified.

Charité *et al.*<sup>2</sup> and Niswander *et al.*<sup>3</sup> have tackled the central issue of which of these phenomena are causal in the chain of events leading to a limb and which are consequential. That is, which genes establish the ZPA and which respond to it? Experiments with Hox genes have, until now, been rather disappointing. Hox gene knockouts in mice have produced minor alterations to carpal bones<sup>4</sup>, a reduction in size of the digits<sup>5</sup> or had no effect on the limb, and it has been suggested that these genes may not be involved directly in the patterning process but instead determine local growth rates of the cartilage primordia<sup>5</sup>. However, overexpression of *Hoxd-6* did induce an extra digit in the chick wing bud<sup>6</sup>, giving a hint of what might be to come.

Sure enough, Charité *et al.* have now taken things further, using the *Hoxb-8* gene which is normally expressed in the posterior half of the mouse forelimb bud. They constructed a transgene in which *Hoxb-8* was placed under the control of the RA receptor- $\beta$ 2 promoter. The gene was then expressed in a mirror-image fashion in some of the forelimb buds, leading to the appearance of an ectopic *Shh* domain on the anterior margin of the limb bud and the induction of FGF4 expression in the anterior AER. The net result was remarkable mirror-image duplications of the forelimbs, the first of their kind in a mammal. The hindlimbs of the transgenic mice were normal, demonstrating that somewhere in the Hox code there must be something that determines the difference between fore- and hindlimbs. Whether this dramatic result was due to the choice of gene cluster or the choice of promoter remains a matter of conjecture.

This result shows that Hox genes are as 'causal' as RA, *Shh* and the ZPA. But which comes before which, or are they all interacting, and where do FGF4 and AER fit in? Niswander *et al.* approach these issues by identifying the role of the AER.



a, How to make a limb: the factors involved. The zone of polarizing activity (ZPA), a group of cells on the posterior edge of the bud which induces a mirror-image set of digits when grafted to the anterior side of another limb bud<sup>8</sup>. Retinoic acid (RA), a metabolite of vitamin A, which is present in higher concentration on the posterior side<sup>9</sup> and also induces duplications<sup>10</sup>. Genes of the *Hoxd* complex, which are expressed in nested sets centred on the posterior side<sup>11,12</sup>; genes of the *Hoxa* complex, expressed in a proximodistal nested set<sup>13</sup>; and other

homeobox-containing genes, generally expressed at the distal tip of the limb bud<sup>14</sup>. The sonic hedgehog gene (*Shh*), the expression pattern of which colocalizes with ZPA and which also induces duplications<sup>15</sup>. Bone morphogenetic proteins (BMP) 2 and 4 (ref. 16). The apical ectodermal ridge (AER), a thickened ridge of ectoderm which is needed to keep the mesenchymal cells dividing and to maintain ZPA activity<sup>17</sup>. Fibroblast growth factor (FGF) 4, which can replace AER function and is expressed in the posterior half of the AER<sup>18</sup>.

The signal(s) for the induction of limb duplication, for example by the ZPA or RA, somehow travel via the AER. Niswander *et al.* first show that grafts of either RA- or *Shh*-transfected cells on the anterior margin of the chick limb bud induce FGF4 in the anterior ridge as well as *Hoxd* expression in the mesoderm. So the sequence of signalling events could be RA → *Shh/Hoxd* → FGF4 in the ridge → outgrowth. But they found that FGF4 was induced by RA before *Shh*. Thus the sequence might be RA → FGF4 → *Shh/Hoxd* → outgrowth. This was confirmed by showing that RA does not induce *Shh* or *Hoxd* without a ridge, but if both RA and FGF4 are applied without a ridge then *Shh* and *Hoxd* are induced. However, this linear sequence of events is not strictly the case because FGF4 alone does not induce *Shh*.

The sequence must therefore be RA + FGF4 → *Shh/Hoxd* → outgrowth. But it also emerged that FGF4 alone can maintain *Shh* once it has been established. So there is a feedback loop between *Shh* and FGF4 in the ridge which obviates any further need for RA once *Shh* has been established. This feedback loop between ectoderm (FGF4) and mesoderm (RA/*Shh/Hox*) must be one of the best characterized examples of epithelial/mesenchymal interactions in molecular terms.

The upshot of this new work is that a specific research agenda can be formulated. The questions to be addressed are these. What is the relationship between the genes of the *Hoxa*, *Hoxb* and *Hoxd* clusters? Does *Hoxb-8* turn on the nested set of *Hoxd* genes which then turn on the nested set of *Hoxa* genes, or do they all feedback on each other? When does RA normally function? Does it establish the domain of expression of *Hoxb-8* during

the establishment of the primary body axis<sup>7</sup>? Does *Shh* act via FGF4 when inducing the *Hoxd* genes? Do the *Hoxd* genes in turn act via FGF4? What happens to the limbs of a *Hoxb-8* or *Shh* knockout mouse? Can an embryo develop limbs without any RA? The answers to even some of these questions will constitute

#### ASTROPHYSICS

## Black holes take centre stage?

Jonathan E. Grindlay

BLACK holes have, increasingly, become central to a range of problems in astronomy, such as the nature of quasars, and have attained a familiarity that belies their difficulty of observation. It has long been suspected that the centre of our Galaxy harbours a monster, a black hole about a million times the mass of the Sun, because the cosmic dictum that we are not special would argue for similarity with what is observed in other galactic nuclei and because the orbits of gas clouds and stars near the centre suggest an unseen central mass. But on page 589 of this issue<sup>1</sup>, Goldwurm *et al.* report that the putative central hole is nowhere to be seen in the first high-sensitivity, hard X-ray imaging survey of the Galactic Centre.

This is disconcerting, as black-hole candidates are increasingly picked out by their characteristic X-ray spectra which extend up to 'hard X-ray' energies above 100 keV. The observation does not rule out a central hole, but requires it to be surprisingly dormant. At the same time, the survey further supports the notion that black holes of a few stellar masses, detected as hard X-ray sources that are usually found in binaries with low-mass companions, may be surprisingly common in the Galaxy.

The story begins with Laplace, who realized that newtonian physics dictated that a large enough mass,  $M$ , in a small enough radius,  $R$ , would yield a gravitational escape velocity  $(2GM/R)^{0.5}$  equal to the speed of light. So black holes have been postulated to be (as the newspapers frequently quote) 'regions of space from which not even light can escape'. Their very blackness would seem to doom them to invisibility. But their compactness ensures that they are such looming gravitational pits as to warp the orbits or gas dynamics of nearby matter.

The first strong black-hole candidate came in 1972 with the identification<sup>2</sup> of the peculiar and highly variable X-ray source Cygnus X-1. This is a massive binary star system in which the unseen compact object emits X-rays as matter from its luminous companion spirals towards it, after being heated to X-ray temperatures of  $10^8$ – $10^9$  K by the viscous dissipation of its

part three of the limb-bud story, and we should not have to wait another 15 years to read it. □

Malcolm Maden is in the Developmental Biology Research Centre, King's College London, 26–29 Drury Lane, London WC2B 5RL, UK.

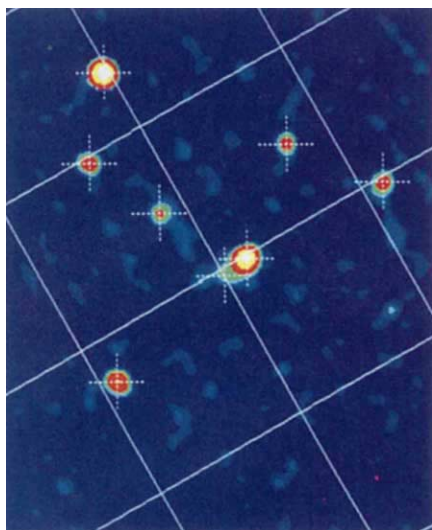
gravitational free-fall energy. Measuring the radial velocity variations of the companion star showed that the compact object (and X-ray source) was most probably at least seven solar masses ( $M_{\odot}$ )<sup>3</sup> and thus above the generally agreed maximum mass of  $3M_{\odot}$  for a neutron star. In other words, the object should have collapsed to a black hole.

Cyg X-1 was the first X-ray binary system identified as a black-hole candidate, but hardly the last, and it is no longer even the most compelling case. Similar work on some half-dozen likely systems<sup>3</sup> shows that in three of these, with surprisingly similar X-ray nova-like outbursts and low-mass companion stars, even the reduced mass probably exceeds the neutron star limit (the reduced mass is the minimum mass for the compact object in the limiting case of zero companion star mass and 90° inclination of the orbital plane, and thus a conservative estimate). With this growing stable of candidates, the most consistent theme is not an X-ray signature of rapid variability but spectacular hard X-ray emission (often out to 100 keV or more). We could identify black-hole candidates, so the idea went, by their hard X-ray emission, distinguishing them from neutron stars (the high-luminosity ones, at least) whose soft thermal black-body emission from their solid surface would cool any hot surrounding corona through the Compton effect<sup>4</sup>.

Now to the Galactic Centre, and active galactic nuclei more generally. Even before the discovery of Cyg X-1 and the unanticipated candidates with low-mass companions, accretion onto super-massive black holes had been suspected as the ultimate power source for the most luminous objects in the Universe: the quasars<sup>5</sup>. The idea was simple: even if such a black hole was not required 'now', it would be later. If the quasar was powered, say, by a massive object supplying rotational energy, or by a cluster supplying nuclear energy in successive supernovae, the emission of so much energy from so small a volume (as indicated by the remarkably rapid variability of quasars) would lead to inevitable collapse<sup>6</sup>. Accretion, and release of gravi-

- Slack, J. *Nature* **279**, 583–584 (1979).
- Charité, J., de Graaff, W., Shen, S. & Deschamps, J. *Cell* **78**, 589–601 (1994).
- Niswander, L., Jeffrey, S., Martin, G. R. & Tickle, C. *Nature* **371**, 609–612 (1994).
- Small, K. M. & Potter, S. S. *Genes Dev.* **7**, 2318–2328 (1993).
- Dolle, P. *et al.* *Cell* **75**, 431–441 (1993).
- Morgan, B. A., Izpisua-Belmonte, J.-C., Duboule, D. & Tabin, C. J. *Nature* **358**, 236–239 (1992).
- Bryant, S. V. & Gardiner, D. M. *Dev. Biol.* **152**, 1–25 (1992).
- Saunders, J. W. & Gasseling, M. T. in *Epithelial-Mesenchymal Interactions* (eds Fleischmajer, R. & Billingham, R. E.) 78–97 (Williams & Wilkins, Baltimore, 1968).
- Thaller, C. & Eichele, G. *Nature* **327**, 625–628 (1987).
- Tickle, C., Alberts, B. M., Lee, J. & Wolpert, L. *Nature* **296**, 564–565 (1982).
- Izpisua-Belmonte, J.-C. *et al.* *Nature* **350**, 585–589 (1991).
- Nohno, T. *et al.* *Cell* **64**, 1197–1205 (1991).
- Yokouchi, Y., Sasaki, H. & Kuriawa, A. *Nature* **353**, 443–445 (1991).
- Izpisua-Belmonte, J.-C. & Duboule, D. *Dev. Biol.* **152**, 26–36 (1992).
- Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Cell* **75**, 1401–1416 (1993).
- Francis, P. H., Richardson, M. K., Brickell, P. M. & Tickle, C. *Development* **120**, 209–218 (1994).
- Summerbell, D. J. *Embryol. exp. Morph.* **32**, 651–660 (1974).
- Niswander, L. *et al.* *Cell* **75**, 579–587 (1993).





Close, but not close enough — X-ray sources<sup>1</sup> near the heart of our Galaxy (gridlines are at 5° intervals).

tational energy, was and is the simpler hypothesis.

The idea seemed to be confirmed most dramatically this May in images of the nucleus of the nearby active radio galaxy, M87, from the restored Hubble Space Telescope<sup>7</sup>. A central gaseous disk could be seen, and the measured gas velocities (again, just proportional to  $(GM/R)^{0.5}$ ) convincingly required the central mass of the nucleus to be about  $2 \times 10^9 M_\odot$ , far in excess of the visible matter and stars but about what had been suspected for a central black hole. Similar measurements of the gas clouds near our own Galactic Centre<sup>8</sup> had long suggested a central (unseen) mass of some  $10^6 M_\odot$ .

So the lack of detectable hard X-ray emission from Sagittarius A\*, the supposed Galactic Centre, strains an already growing problem: why is there no sign of accretion onto the central black hole from the gas clouds (and possibly mass-losing stars in a dense central cluster) observed in the central few parsecs? The long history of X-ray detection of sources coincident with or near Sgr A\* had always pointed to a luminosity problem<sup>8</sup>, as the measured soft X-ray luminosity was no more than about  $10^{35}$  erg s<sup>-1</sup>, or only about  $10^{-7}$  of the Eddington luminosity (if the spectrum were extrapolated to ~100 keV) for a  $10^6 M_\odot$  black hole. But more luminous hard X-ray or soft gamma-ray sources had been seen somewhere near the Galactic Centre with earlier non-imaging detectors and so it was hoped that Sgr A\* was a respectable active galactic nucleus after all.

The new observations from the Sigma/GRANAT telescope<sup>1</sup>, with unprecedented total exposure time of nearly 1,800 hours over 4 years, show that this is not the case. The luminous hard X-ray sources around the Galactic Centre are in fact isolated, discrete black-hole candi-

dates, accreting matter either from binary companions (and one X-ray nova is included) or from dense molecular clouds through which they are drifting. The newly developed coded aperture technique for hard X-ray imaging had already revealed<sup>9</sup>, in earlier balloon-borne hard X-ray images, at least one transient black-hole candidate in the Sgr A\* region, with possible backscattered emission at 511 keV (this is an indication of positron annihilation, which was also probably directly detected by Sigma from a flare of the brightest black-hole candidate<sup>1</sup>). The Galactic Centre region alone suggests that the whole Galaxy sustains a large population of black holes of mass around  $10 M_\odot$  in low-mass binaries (more than  $10^3$ ) or in isolation (more than  $10^7$ ).

The black-hole candidates, with their nova-like extreme variability in X-rays (and in the optical), may point to the answer: the central massive black hole in the Galactic Centre is dormant, with mass not now accreting onto the central hole but gathering in a surrounding accretion disk for some future outburst. The X-ray novae seem to have duty cycles of luminous accretion of only about a month every 50 years or so and may be governed by an accretion limit-cycle instability in their accretion disk<sup>10</sup>. During their 'off' states, their accretion rate (and thus X-ray luminosity) is also about  $10^{-7}$  of their Eddington value. If accretion onto the black hole in Sgr A\* had a similar instability mechanism we should hardly expect to see it.

Thus although the rare class of galaxies with active nuclei may point to super-massive black holes now undergoing accretion, normal galaxies like the Milky Way or Andromeda (M31) may also harbour central massive black holes (although with masses lower by factors of  $10$ – $10^3$ ). The real difference is not the mass, or the mere presence; rather, it is the phase of the duty cycle. The active galactic nuclei are simply in the relatively rare 'on' state, while the normal galactic nuclei, such as Sgr A\*, are dormant like the vast majority of their brethren. □

Jonathan E. Grindlay is at the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

1. Goldwurm, A. *et al.* *Nature* **371**, 589–591 (1994).
2. Bolton, C. T. *Nature* **235**, 271–273 (1972).
3. Van Paradijs, J. & McClintock, J. E. in *X-ray Binaries* (eds Lewin, W. H. G., van Paradijs, J. & van den Heuvel, E. P. J.) (Cambridge Univ. Press, in the press).
4. Sunyaev, R. *et al.* *Astr. Astrophys.* **247**, L29–L32 (1991).
5. Zeldovich, Y. B. & Novikov, I. D. *Dokl. Acad. Nauk. SSSR* **158**, 811 (1964).
6. Rees, M. J. *Observatory* **98**, 210–232 (1978).
7. Ford, H. *et al.* *Astrophys. J. Lett.* (in the press).
8. Genzel, R., Hollenbach, D. & Townes, C. *Rep. Prog. Phys.* (in the press).
9. Grindlay, J. E., Manandhar, R. & Covault, C. *Astr. Astrophys. Suppl.* **97**, 155–158 (1993).
10. Minishige, S., Kim, S. & Wheeler, J. C. *Astrophys. J. Lett.* **358**, L5–L8 (1990).

DAEDALUS

## Instant Newzak

RADIO news has many advantages over the television variety. It is not burdened with compulsory pointless pictures; it doesn't seem obliged to make its newsreaders into celebrities; and you can catch up with it while doing other things. It could be brief, efficient and matter-of-fact. There was a blessed time when BBC radio could straightforwardly announce, "There is no news tonight." And yet nowadays radio news expands to fill the space available for it — whole channels of desperate waffle in some cases. Daedalus wants to prune it back to its original efficiency.

In the studio, news exists as a number of stories in a 'stack', whose order depends on their newness and perceived importance. Any item slides down the stack with time; it features ever later and more briefly in successive newscasts, until it falls off the stack entirely. The simplest way of getting the stack to its consumers would simply be to broadcast it as a succession of updates, and accumulate it at the receiver for playback at any time.

The technology looks very simple. A sparse traffic of audio updates would not need a channel of its own; it could be coded and spread out to ride in the spaces of another radio channel, as Teletext does on a television transmission. At the receiver, it would be decoded and held in a memory chip of the type used in voice-mail and some audio processors. The receiver would update the news stack continuously as a 'background activity' — even when nominally switched off. The listener could hear the current stack at any time merely by touching a button. Compared with the money now being spent to realize the imagined need for video-on-demand transmissions, this news-on-demand system would be cheap indeed.

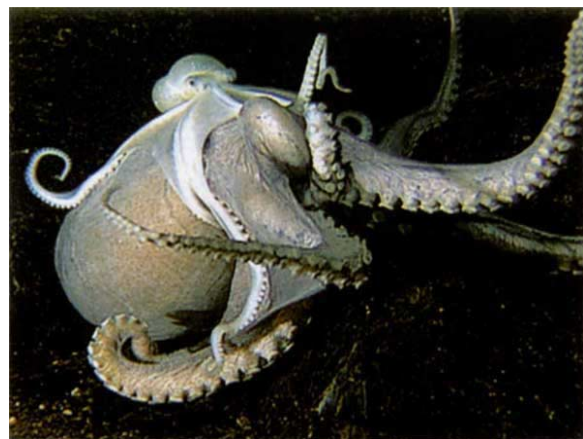
Furthermore, the listener could program and customize it to suit himself. A subject code attached to each item could be read by a programmable filter on the receiver itself. The listener could set the program to play back just those items he cared about (sport, politics, or whatever) and ignore the rest. News fanciers could stay abreast of their various fields of interest without having to sit through hours of irrelevant waffle. Television news (whose useful information is nearly all carried by its sound anyway) might lose some of its dangerous hypnotic power; millions of viewing hours could be saved for reality. And those of us who attend to the news purely negatively, to reassure ourselves that nothing important has happened, could gain that reassurance on demand, at any time.

David Jones

# Close encounter in the deep

**SIR** — Benthic octopods are cosmopolitan marine predators that have been reported to occur to depths of almost 4,000 m (ref. 1), yet *in situ* observations of the animals in deep-sea environments have been extraordinarily rare. In Decem-

ber 1993, the deep-diving submersible *Alvin*, while traversing expansive lava sheet flows on the floor of the axial summit caldera of the East Pacific Rise at 9°50'N (depth 2,512 m)<sup>2</sup>, encountered two males of distinct octopodid species engaged in copulatory behaviour (see figure). Behavioural patterns of the two octopods were documented over a 16-minute period using a high-resolution, 3-chip camera; the Beta videotape record is the basis for observations reported here.



During the 16-minute video sequence, the small male shifted from his initial transverse orientation to a more parallel position relative to the large male. When observations began, the hectocotylus of the small male was near the left side of the large male's mantle aperture. As the small male shifted position, his hectocotylus moved across the posterior dorsal surface of the large male's mantle; his ligula (intromittent organ), in contrast to its normal, protected position, became extended. After the shift in position, his hectocotylus was brought anteriorly to the large male's right mantle aperture. The

small male subsequently succeeded in introducing the distal end of his hectocotylus into the mantle cavity of the large male (see figure), an act constituting copulation in octopodids<sup>3</sup>. Ventilation rate, which is argued to reflect alteration of behavioural states in octopodids<sup>4</sup>, was monitored in the small male throughout the 16-minute video sequence. During the initial 10 minutes, ventilation by this individual was virtually undetectable. During the last 6 minutes of observation, the small male vigorously ventilated his mantle 7–9 times each minute. It was during the latter period that his ligula became extended and his hectocotylus entered the mantle aperture of the large male.

The present study, to our knowledge the first account of copulatory behaviour in deep-water octopods, questions whether copulation between males of different species of these mobile abyssal organisms is a rare event. The apparent readiness to copulate, or attempt to copulate, may reflect adaptations to low mate availability and the short lifespan of cephalopods, a group in which promiscuous mating occurs<sup>3</sup>. Although females can increase their fitness by mating with older males (those with a proven ability to survive to an advanced age)<sup>5</sup>, males are fully mature for a relatively brief period before senescence<sup>3</sup>. The fitness of mature males carrying accrued spermatophores is limited by the time remaining in their lifespans and by the number of mates they encounter. Among shallow-water octopuses, mature males are more mobile than immature individuals, increasing their likelihood of encountering mates<sup>6</sup>. Where mates are few, such as in the low-density, heavily fragmented populations of the deep sea, high mobility may not be sufficient to increase male fitness. In such environments, males may further modify their behaviour to increase the time, energy and possibly gametes

invested in each encounter with another octopod rather than to leave opportunities for reproduction unexplored.

**Richard A. Lutz**

Department of Marine and Coastal Sciences,

Rutgers University,

New Brunswick, New Jersey 08903, USA

**Janet R. Voight**

Department of Zoology,

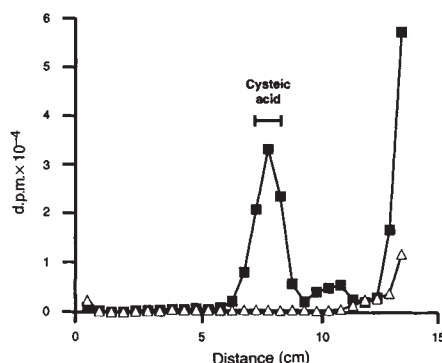
The Field Museum of Natural History,

Roosevelt Road at Lake Shore Drive,

Chicago, Illinois 60605, USA

## Preventing hair loss in mice

**SIR** — We wish to report the successful transfer to mice of two genes, *cysE* and *cysK*, from *Escherichia coli* which encode the enzymes necessary for the assimilation of inorganic sulphide for the biosynthesis of the amino-acid cysteine<sup>1,2</sup>. These mice are able to synthesize cysteine when supplemented with Na<sub>2</sub>S, and the effect of dietary sulphur amino-acid deprivation on hair growth is prevented. Our results demonstrate that the introduced biochemical pathway is functional and



**FIG. 1** Biosynthesis of cysteine (measured as cysteic acid) from Na<sub>2</sub><sup>35</sup>S *in vitro* in intestinal tissue isolated from transgenic mice containing the gene *MTCEK1*. Transgenic, filled squares; control, triangles.

establishes the principle implicit in the pioneering research of Palmiter *et al.*<sup>3,4</sup>; that the modification of animal biochemistry is feasible by the transfer of functional genes from bacteria or other sources in nature.

- Voss, G. L. *Malacologia* **29**, 295–307 (1988).
- Haymon, R. *et al.* *Earth planet. Sci. Lett.* **104**, 513–534 (1991); **119**, 85–101 (1993).
- Mangold, K. in *Cephalopod Life Cycles* Vol. 2 (ed. Boyle, P. R.) 157–200 (Academic, London, 1988).
- Boyle, P. R. *J. exp. mar. Biol. Ecol.* **69**, 129–136 (1983).
- Halliday, T. R. in *Mate Choice* (ed. Bateson, P.) 3–32 (Cambridge Univ. Press, UK, 1983).
- Voight, J. R. *J. Zool., Lond.* **228**, 247–263 (1992).

- Denk, D. & Bock, A. *J. gen. Microbiol.* **133**, 515–525 (1987).
- Byrne, C. R., Monroe, R. S., Ward, K. A. & Kredich, N. M. *J. Bact.* **170**, 3150–3157 (1988).
- Palmiter, R. D. *et al.* *Nature* **300**, 611–615 (1982).
- Palmiter, R. D., Norstedt, R. E., Gelinas, R. E., Hammer, R. E. & Brinster, R. L. *Science* **222**, 809–814 (1983).
- Leish, Z., Byrne, C. R., Hunt, C. L. & Ward, K. A. *Appl. Environ. Microbiol.* **59**, 892–898 (1993).
- Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo. A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1986).
- Huovinen, J. A. & Gustafsson, B. E. *Biochim. biophys. Acta* **136**, 441–447 (1967).
- Leveille, G. A., Sauberlich, H. E. & Shockley, J. W. *J. Nutr.* **75**, 455–458 (1961).



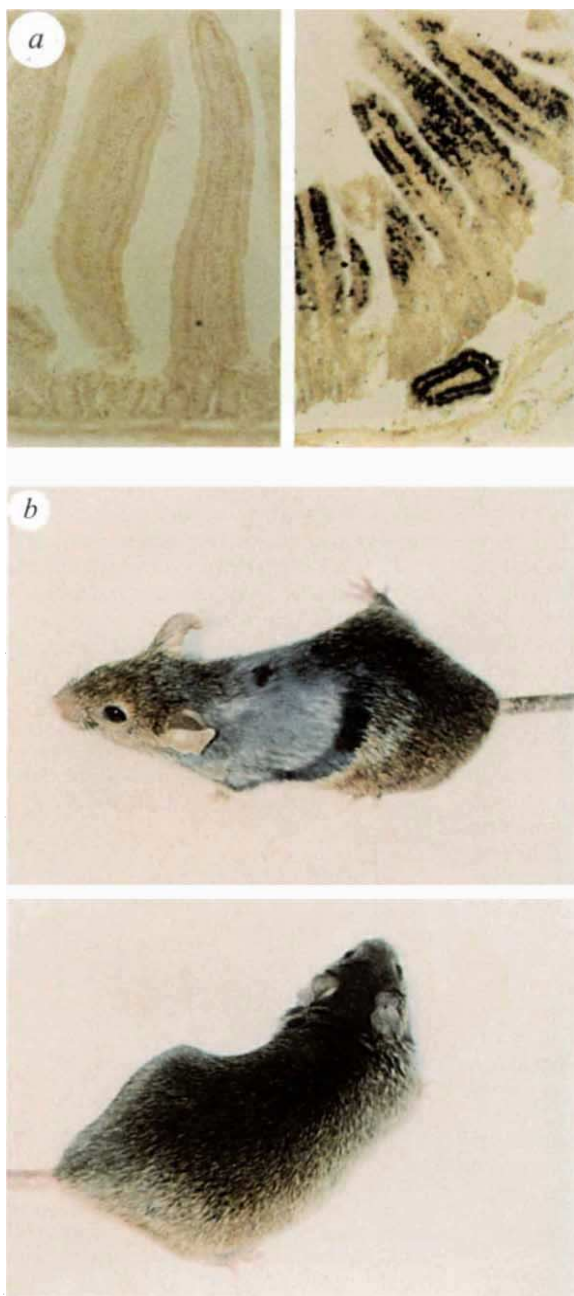


FIG. 2 *a*, *In situ* analysis of the expression of the gene *MTCEK1* in intestinal epithelia of non-transgenic (left) and transgenic (right) mice. *b*, Protective effect of the cysteine biosynthetic pathway on diet-induced hair loss in mice. Age-matched non-transgenic (top) and transgenic (bottom) mice were placed on a synthetic diet in which the combined sulphur amino acids were reduced to 0.1% (w/w) of the total diet, but supplemented with 0.12% (w/w)  $\text{Na}_2\text{S}$ . Throughout the experiment, the drinking water of the mice was supplemented with 25 mM  $\text{ZnSO}_4$ . Photographs were taken after 7 days. Hair loss was observed in the non-transgenic mice but not in the transgenic animals.

The gene *pMTCEK1* (ref. 5), which encodes the enzymes serine acetyl transferase and *O*-acetylserine sulphydrylase, was transferred to transgenic mice by conventional microinjection<sup>6</sup>, breeding lines were established and expression of the transgene was demonstrated by northern blots and enzyme analyses. The pre-

sence of a functional biosynthetic pathway was confirmed by incubating dissected intestinal tissue with radioactive  $\text{Na}_2\text{S}$  and examining the synthesized amino acids. The results demonstrate a substantial peak of radioactivity in transgenic tissue incubations in the position at which cysteine (measured as cysteic acid) migrates (Fig. 1).

The expression of the *MTCEK1* gene in the intestinal epithelium of the transgenic mice was verified by an *in situ* analysis using an antisense bacterial *cysE* sequence RNA probe (Fig. 2*a*). The *cysE* mRNA could only be detected in the cells of intestinal epithelia from the transgenic mice.

The operation of the new biochemical pathway *in vivo* was demonstrated by placing the transgenic and control mice on a synthetic diet deficient in sulphur-containing amino acids but supplemented with  $\text{Na}_2\text{S}$  (refs 7, 8). Control mice showed either an extensive hair loss on the back and sides (three mice) or a marked 'dishevelled' quality in their coats (nine mice). On the other hand, all the transgenic mice (14) retained a normal hair coat with no sign of hair loss or change in coat quality. A comparison of a transgenic and a control mouse is shown in Fig. 2*b*. Because the hair loss in the controls could be reversed by supplementation of the diet with sulphur amino acids, we interpret our observations as indicating that the operation of the cysteine biosynthetic pathway in the transgenic mice has provided these animals with sufficient cysteine to prevent the diet-induced hair loss.

These results demonstrate that a functional biochemical pathway can be transferred from bacteria to mammals, can use endogenous substrates and co-factors for its effective operation and is capable of preventing some of the phenotypic effects associated with specific nutrient deprivation.

Further details of these results are

available on request from the authors and will be published in detail elsewhere.

**Kevin A. Ward, Zdenka Leish, John Bonsing, Noriko Nishimura, Graham R. Cam, Alan G. Brownlee, Colin D. Nancarrow**  
Division of Animal Production,  
CSIRO,  
Locked Bag 1,  
Delivery Centre, Blacktown,  
New South Wales, Australia 2148

## Smokeless fire

**SIR** — Summerlin<sup>1</sup> objects, in Scientific Correspondence, to the suggestion by Rodu and Cole<sup>2</sup> that smokers switch to using smokeless tobacco as a habit that would be significantly less harmful to them than active smoking. Summerlin bases his objection on the risk to users.

But smokers are not the only patients to whom health-care providers have responsibilities. Wells, in the most recent quantification of deaths due to second-hand smoking<sup>3</sup>, suggests "that in 1985 an estimated 62,000 ischemic heart disease deaths in the United States were associated with exposure to environmental tobacco smoke". He had previously estimated<sup>4</sup> that in the same year 3,000 lung-cancer patients and 11,000 patients with "other cancers" died in the United States because of second-hand smoking. That is a grand total of 76,000 estimated deaths, rather than 30,000 estimated cases of oral cancer.

A massive shift towards the use of oral tobacco by smokers would considerably cut the tens of thousands of deaths caused by tobacco-smoke pollution. On balance, it would seem that health-care providers owe more to their non-smoking patients than is owed to active smokers who might choose to switch to oral tobacco.

**Phillip Whidden**  
Association for Nonsmokers' Rights,  
Melgund Terrace,  
Edinburgh EH7 4BU, UK

1. Summerlin, D-J. *Nature* **371**, 113 (1994).
2. Rodu, B. & Cole P. *Nature* **370**, 184 (1994).
3. Wells, A. J. *J. Am. Coll. Cardiol.* **24**, 546-554 (1994).
4. Wells, A. J. *Environ. Int.* **14**, 249-265 (1988).

### Scientific Correspondence

Scientific Correspondence is a relatively informal section of *Nature* in which matters of general scientific interest, not necessarily those arising from papers appearing in *Nature*, are published. Because there is space to print only a small proportion of the letters received, priority is usually given according to general interest and topicality, to contributions of fewer than 500 words, and to contributions using simple language.

## Biodiversity in model ecosystems

**STR** — In his News and Views article, Kareiva<sup>1</sup> applauded the innovative approach of Naeem *et al.*<sup>2</sup>, who carried out experiments on ecosystems under controlled and duplicable conditions. Kareiva is of the view that field experiments on ecosystems will always be necessary, but they tend to be hard to repeat. This has often been a stumbling block in ecology research. Experiments on small microcosms do not suffer from this difficulty.

One criticism of microcosm-based research is that enclosure in a small volume is somehow 'unnatural'. But our terrestrial biosphere, the best natural example, is essentially a closed system. Enclosure (not necessarily full closure) paves the way to an experimental approach in ecology that is more receptive to advanced instrumentation, and consequently offers a better control over experimental conditions, and allows quantification of the observations and a powerful exploration of their validity by means of comparison and repetition.

In the experiment designed by Naeem *et al.*<sup>2</sup>, the productivity of three ecosystem communities (their ability to produce biomass) is shown to rank parallel to their biodiversity. None of the other ecosystem processes measured (decomposition, nutrient retention and water retention) exhibited a consistent pattern of variation among treatments. We believe that the choice of plant species in these three communities is not entirely justified in that it does not appear to have been motivated by a concern about biodiversity loss. We question the selection of *Senecio vulgaris* and *Stellaria media* in the poorest community instead of, for example, larger species in the pool considered. Also, there is a mismatch between the size of the large plants (about 60 cm) and the small size of the soil samples (1 m × 1 m) to be colonized. Attempts to balance the three communities based on the initial number of seeds (80 seeds for each community) is a necessary, but perhaps not sufficient, precaution. In particular, large plants (taller than 50 cm) were present only in the medium- and high-diversity communities whereas the low-diversity community consisted only of small plants (smaller than 30 cm). It is well known that a monospecific culture can be as productive as a mixed culture, provided that the species has been selected, or adapted, to exhibit this potential. Another choice for the first community, featuring a large plant species with a canopy structure capable of optimizing the three-dimensional colonization of the available space, might have yielded a distinct ranking.

We appreciate that the purpose of the

study of Naeem *et al.* was not to identify the surviving species in an environment where biodiversity is declining, but we wish to draw attention to the difficult problem of adequately 'selecting' species to simulate a low-diversity community. Provided special attention is devoted to this problem, similar experimental approaches, in which producer-consumer-decomposer interactions in controlled environments are integrated, offer a new and promising context for studying biodiversity in liaison to ecosystem stability. The study of Naeem *et al.* is a pioneering experiment in pointing the way in this direction.

**Marcel André**

**François Bréchinac**

**Pierre Thibault**

CEA-Direction des Sciences du Vivant,

Département Physiologie Végétale et

Ecosystèmes,

Centre d'Etudes de Cadarache,

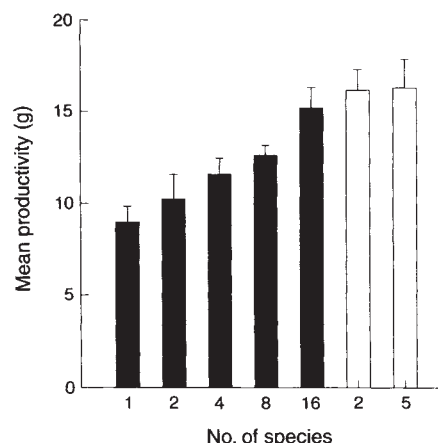
BP 1, 13108-Saint Paul lez Durance Cedex, France

**NAEEM ET AL. REPLY** — André *et al.* are correct to say that the choice of species in a manipulation of diversity critically determines its outcome. They suggest that exclusion of a single plant species, *Chenopodium album*, from the low-diversity treatment in our experiment<sup>2</sup>, may have been solely responsible for the positive association between biodiversity and ecosystem productivity. Species diversity is a difficult variable to manipulate experimentally because of the many possible species combinations such manipulations can generate. We were aware of the importance of the problem and took the following three steps to ensure that our experiment would not be overly sensitive to the presence or absence of one or a few species.

(1) Design. We used relatively ecologically homogeneous species within trophic classes to ensure against the possibility of 'keystone' species effects<sup>1,2</sup>.

(2) Calibration of the Ecotron plant communities. A glass-house experiment (S. N. *et al.*, submitted; see figure) compared the potential productivity of the plant species combinations grown in the Ecotron with the productivities of alternative combinations. Results support findings in the Ecotron in two ways. First, the relationship between productivity and species richness (black bars) is positive. Second, the specific combinations used in the Ecotron (white bars and the black 16-species bar) were equal in potential productivity (ANOVA; d.f. = 2, 25;  $F = 0.23$ ;  $P = 0.79$ ). Thus, the combinations of species used in the Ecotron provided a conservative test of the association between diversity and productivity.

(3) Keystone species effects. *C. album*, although tall, was not a keystone species. Under glass-house conditions, *C. album* is



Calibration of the productivity of Ecotron plant species combinations against other possible combinations (S. N. *et al.*, submitted). Black bars are means of above-ground plant biomass (dry weight  $\pm 1$  s.e.) from pots containing 16 individual plants that ranged in diversity between 1 (monocultures), 2, 4, 8, (intermediate richness combinations) and 16 species (full richness combinations) from the pool of 16 plant species used in the Ecotron experiment<sup>2</sup>. Intermediate richness combinations were random selections of species, without duplication, drawn from the pool of 16. White bars, the specific 2-species and 5-species combinations used in the Ecotron. The number of replicates were 4 for each monoculture and 10 for full richness combinations. Intermediate richness combinations consisted of 19, 30 and 40 unique combinations of 2, 4 and 8 species, respectively. Ecotron combinations of 2 species and 5 species had 10 and 8 replicates, respectively. Soil was identical to soil used in the Ecotron. Plants were grown between 10 May and 16 July 1993 (one plant generation) in a glass house at Silwood Park.

not the most productive species, ranking sixth in productivity; *Senecio vulgaris* and *Stellaria media*, the plant species universal to all treatments, ranked second and third highest in productivity. Second, *C. album* accounted for 20.4% of sampling pin encounters with vegetation in intermediate- and 8.36% of encounters in high-diversity mesocosms. These measured abundances accord with expected encounters of 20.0% ( $1/5$  of species) and 6.25% ( $1/16$  of species) for intermediate- and high-diversity mesocosms, respectively.

**Shahid Naeem\***

**Lindsey J. Thompson**

**Sharon P. Lawler\***

**John H. Lawton**

**Richard M. Woodfin**

NERC Centre for Population Biology,

Imperial College at Silwood Park,

Ascot, Berks SL5 7PY, UK

\*Present addresses: Department of Ecology, Evolution and Behaviour, University of Minnesota, St Paul, Minnesota 55108, USA (S.N.); Morgan School of Biological Sciences, University of Kentucky, Lexington, Kentucky 40506, USA (S.P.L.).

1. Kareiva, P. *Nature* **368**, 686–687 (1994).

2. Naeem, S. *et al.* *Nature* **368**, 734–737 (1994).



# Terrestrial carbon storage at the LGM

**SIR** — Several studies have yielded discrepant estimates for the change in carbon storage from the Last Glacial Maximum (LGM) to pre-industrial times which range widely from ~0 to +1,350 Pg (1 Pg=10<sup>15</sup> g) (refs 1–6).

Two recent advances now permit a more rigorous evaluation of the available estimates of change in terrestrial carbon storage using the change in whole ocean  $\delta^{13}\text{C}$  value from LGM to pre-industrial times<sup>6</sup>: (1) the  $\delta^{13}\text{C}$  value of atmospheric carbon dioxide is known to have been 0.3–0.7‰ lower than present during the LGM (see ref. 7); and (2) it is possible to estimate the bulk  $\delta^{13}\text{C}$  value of carbon in pre-industrial biomass. From this information it is possible to calculate what the  $\delta^{13}\text{C}$  value of terrestrial biomass at the LGM would have to be in order to satisfy the mass/isotope balance for a specified change in carbon storage (see figure).

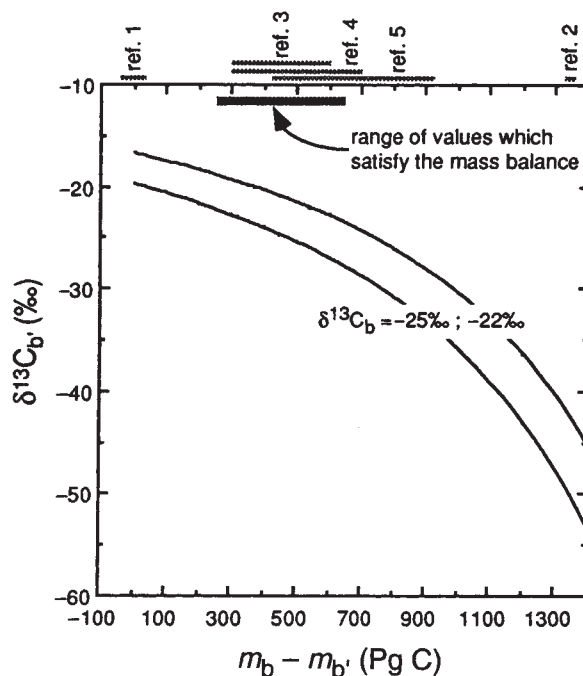
The  $\delta^{13}\text{C}$  value of carbon in pre-industrial terrestrial biomass can be estimated in two ways. The first is based on over 250 carbon-isotope published<sup>8</sup> and unpublished analyses of surface soils from C3 biomes spanning latitudes from 0–60° and 4,600 m of altitude. The  $\delta^{13}\text{C}$  values of surface soils in those biomes in which C4 species are present are more difficult to constrain, but can be specified within broad limits from published soil  $\delta^{13}\text{C}$  values. This approach suggests that the isotope composition of carbon in the pre-industrial terrestrial biosphere ( $\delta^{13}\text{C}_b$ ) had a value of not less than –25‰, using the range of available estimates for modern carbon storage in each of the major biomes<sup>1–5</sup>. An alternative modelling approach based on physiological considerations<sup>9</sup> yields a bulk  $\delta^{13}\text{C}$  value of not higher than –22‰ for carbon in pre-industrial terrestrial biomass.

The curves in the figure depict the relationship between change in carbon storage from the LGM to pre-industrial times ( $m_b - m_{b'}$ ) and the carbon-isotope composition required for carbon in terrestrial biomass at the LGM ( $\delta^{13}\text{C}_b$ ) to balance a whole-ocean  $\delta^{13}\text{C}$  value of –0.32‰ (for  $\delta^{13}\text{C}_b = -22$  to –25‰). It is immediately apparent that the high estimate of Adams *et al.*<sup>2</sup> (1,350 Pg) is implausible as it would require that the  $\delta^{13}\text{C}$  value of the terrestrial biosphere be ~–50‰ at the LGM. The remaining published estimates yield more reasonable values, so to narrow the field further it is necessary to make an estimate of the difference in  $\delta^{13}\text{C}$  value of carbon in terrestrial biomass between LGM and pre-industrial times.

Using the estimates for  $\delta^{13}\text{C}_b$  (discussed above), the  $\delta^{13}\text{C}$  value of the LGM atmosphere and incorporating the possibility that isotope discrimination in C3 plants was reduced at the LGM<sup>10</sup>, it is possible to estimate the difference in the bulk  $\delta^{13}\text{C}$  value of carbon in the terrestrial biosphere at the LGM relative to pre-industrial times using the published biome distributions at the LGM<sup>1–5</sup>. Such a calculation suggests that the bulk  $\delta^{13}\text{C}$  value of carbon in the terrestrial biosphere at the LGM was 0–2‰ enriched in <sup>13</sup>C relative to the pre-industrial biosphere.

The range of changes in carbon storage which satisfy the mass balance considerations is 310–550 Pg for a  $\delta^{13}\text{C}_b$  of –25‰ to –22‰. From these estimates it can be seen that the low estimate of Prentice and Fung<sup>1</sup> cannot be correct, as their estimate would require a pre-industrial versus LGM difference in the  $\delta^{13}\text{C}$  value of terrestrial carbon of over 5‰ to satisfy the mass balance. If it is assumed that the true value of  $\delta^{13}\text{C}_b$  is not precisely known to be –0.32‰, but can be expected to lie between –0.3 and –0.4‰, then the range of permissible changes in carbon storage increases to 270–720 Pg, still precluding the highest and lowest published estimates. This range is similar to that of Prentice *et al.*<sup>5</sup> and Friedlingstein *et al.*<sup>3</sup>, and overlaps the lower end of the range cited by Van Campo *et al.*<sup>4</sup>.

The 'true' change in terrestrial carbon storage probably lies towards the lower end of the range given above, because (1) C4 species are at a competitive advantage at low  $p\text{CO}_2$  values and therefore may



Calculated  $\delta^{13}\text{C}$  value of carbon in the terrestrial biosphere at the LGM ( $\delta_{b'}$ ) required to satisfy the mass balance for a specified change in carbon storage ( $m_b - m_{b'}$ ), given by the following equation:

$$\delta_{b'} = [\delta_a m_a + \delta_b m_b + \delta_o m_o - \delta_{a'} m_{a'} - \delta_{o'} (m_o + (m_a - m_{a'}) + (m_b - m_{b'}))] / m_{b'}$$

where:  $\delta_a$  =  $\delta^{13}\text{C}$  value of carbon dioxide in pre-industrial atmosphere (–6.5‰);  $m_a$  = mass of carbon in pre-industrial atmosphere (600 Pg);  $\delta_b$  =  $\delta^{13}\text{C}$  value of carbon in pre-industrial terrestrial biosphere (–22 to –25‰);  $m_b$  = mass of carbon in pre-industrial terrestrial biosphere (2,200 ± 100 Pg);  $\delta_o$  =  $\delta^{13}\text{C}$  value of pre-industrial oceanic carbon reservoir (0‰);  $m_o$  = mass of carbon in pre-industrial oceanic carbon reservoir (38,000 Pg);  $\delta_{b'}$  =  $\delta^{13}\text{C}$  value of carbon in LGM biosphere (0 to 2‰ higher than  $\delta_b$ );  $m_{b'}$  = mass of carbon in LGM terrestrial biosphere;  $\delta_{a'}$  =  $\delta^{13}\text{C}$  value of carbon dioxide in LGM atmosphere (–7.0 ± 0.2‰);  $m_{a'}$  = mass of carbon in LGM atmosphere (430 Pg);  $\delta_{o'}$  =  $\delta^{13}\text{C}$  value of oceanic carbon reservoir at LGM (–0.32‰);  $m_{o'} = m_o + (m_a - m_{a'}) + (m_b - m_{b'})$  = mass of oceanic carbon reservoir at LGM. Note: the range for ref. 3 includes 100–300 Pg peat accumulation<sup>5</sup>.

have made up a larger proportion of biomass in any given biome at the LGM<sup>11</sup>; and (2) the mass balance approach used here does not include the possibility that low-<sup>13</sup>C carbon of marine origin may have been sequestered on the continental shelves during post-glacial sea-level rise<sup>6</sup>. Nevertheless, the isotope mass balance approach applied here does go a considerable way to narrowing the range of plausible estimates for changes in carbon storage from the LGM to the present, and is in substantial agreement with several of the previously published estimates.

**Michael I. Bird**

*Research School of Earth Sciences,*

**Jon Lloyd**

*Research School of Biological Sciences,*

*Australian National University,  
Canberra, ACT 0200,  
Australia*

1. Prentice, K. C. & Fung, I. Y. *Nature* **346**, 48–51 (1990).
2. Adams, J. M. *et al.* *Nature* **348**, 711–714 (1990).
3. Friedlingstein, P. *et al.* *Geophys. Res. Lett.* **19**, 897–900 (1992).
4. Van Campo, E., Guiot, J. & Peng, C. *Global Planet. Change* **8**, 189–201 (1993).
5. Prentice, I. C. *et al.* *Global Ecol. Biogeogr. Lett.* (in the press).
6. Crowley, T. J. *Nature* **352**, 575–576 (1991).
7. Leuenberger, M., Siegenthaler, U. & Langway, C. C. *Nature* **357**, 448–490 (1992).
8. Bird, M. I., Haberle, S. & Chivas, A. R. *Global Biogeochem. Cycles* **8**, 13–22 (1994).
9. Lloyd, J. J. & Farquhar, G. D. *Oecologia* (in the press).
10. Van de Water, P. K. *et al.* *Science* **264**, 239–243 (1994).
11. Cole, D. R. & Monger, H. C. *Nature* **368**, 533–536 (1994).

# The political atom

William T. Golden

**The Myths of August.** By Stewart Udall. *Pantheon*: 1994. Pp. 399. \$25.

CONTROVERSY has long surrounded the use by the United States of fission atomic bombs against Japan and the subsequent US development of the much more potent fusion weapon, the hydrogen bomb. As these epochal events recede into the past, revisionist historians who were not there at the time spring up to tell us what the *dramatis personae* should have done when the political climate was very different from what it has become. Such a one is the prominent Stewart Udall. The title of his book refers to the use of atomic bombs against Hiroshima and Nagasaki in August 1945. More broadly, he writes critically of the entire atomic era from the Second World War through the Cold War to today, and he goes on to prescribe a path of conduct for future leaders of the United States.

This arresting but unreliable diatribe against the establishment will attract attention. The bulk of it is simplistic, prejudiced, categorical, one-sided and intolerant. Creating myths of its own, it will gratify true believers, irritate many of those responsible for decisions in times of stress and stimulate thought about complex issues by concerned laymen as well as by open-minded scientists and administrators. The book is easy reading and, although it lacks intentional humour, for some readers a myth is as good as a smile. The author attacks other people's dogmas and substitutes his own. "Let dogma eat dogma", as Herman Kahn put it years ago.

Born in 1920 in Arizona, Udall has had a distinguished and diversified career in private life and in government: as an environmentalist, Congressman for six years and Secretary of the Interior for eight years under Presidents John F. Kennedy and Lyndon Johnson. Udall is an honorable man. Clearly, he is a man of high principle and fervently held beliefs. For many years now, he has been in private practice as a 'personal injury lawyer'.

## Criticism

This book is a plaintiff's brief against most of the men (there were no women) responsible for the US government's atomic and related military policies during and after the Second World War. The author criticizes in severe, pungent and sometimes intemperate terms presidents, cabinet members, the Congress and the judiciary (which held against his clients in two major cases), as well as the military and many scientists. Among the objects of

his condemnation, derision or ridicule are Presidents Franklin D. Roosevelt, Harry S. Truman, Jimmy Carter, Ronald Reagan and George Bush; Secretaries of State Henry L. Stimson, John Foster Dulles, Dean Acheson and James Byrnes; public servants Robert Lovett, Allen Dulles and Paul Nitze; journalist William L. Laurence; Generals Leslie Groves and Curtis Le May; and scientists J. Robert Oppenheimer, Glenn Seaborg, Ernest O. Lawrence, Willard Libby, Harold Urey, Edward Teller and John von Neumann. Praise is bestowed sparingly, notably on Andrei Sakharov and Mikhail Gorbachev, as men of peace, and on David Lilienthal, Merrill Eisenbud and Robert Frost.

There is a fatal underlying flaw in Udall's brief. To support his opinions half a century after the events, he relies on hindsight, hindsight frequently peppered with astigmatic distortions. For example, he says: "Impelled by Olympian hubris, the AEC [Atomic Energy Commission]'s headlong fission power program was marked for failure from its inception." And, because of the policy of secrecy, "it was foreseeable from the outset that officials of the AEC would be tempted to lie, to impede justice, and to disregard duties they owed to unknowing citizens".

The last shall be first. Start with Udall's ten-page final chapter, "Lessons of the Cold War", then see his two-page preface. These set forth his ideas for bettering the world in the decades ahead, now that the Cold War has ended: in brief, return to the ethics and ideals of Abraham Lincoln and of Thomas Jefferson and the other founders of the Republic. Other ingredients of his prescription include expansion of the United Nations (UN) peacekeeping function; reduction of the world's nuclear arsenal, ultimately to zero, through US and UN leadership; dismantling "of the national security state we created to combat Communism"; "drastically reducing the classification of information in order to end the secrecy that has encouraged mendacity in high places"; "ending the clandestine influence of nuclear 'high priests', 'strategy elite', and assorted 'wise men'"; "curbing the authority of the nation's 'security' organizations (the FBI, the CIA and their acolyte agencies)"; and a return to the "Constitution, the Bill of Rights, and the other founding documents of our Republic" and toward "Lincolanian magnanimity". And he concludes with the indictment that "during the sad history of the atomic age and

the Cold War, our political institutions have not failed us, our leaders have betrayed those institutions and thus the American people". The author's approach is more idealistic than realistic. His objective of a more peaceful world is worthy. But, the plight of the underclass in the United States and throughout the world is not addressed; nor is the daunting objective of altering human nature from its Cain-slew-Abel beginnings in order to defuse persistent, ancient ethnic strife and to outweigh the successful power of Darwinian competition.

## Leadership

The proposal by Udall, and many others in the post Cold War world, to strengthen the UN peacekeeping and fire-fighting roles is timely and appealing but not new. A. E. Houseman's "Epitaph on an Army of Mercenaries" is apposite: "What God abandoned, these defended/And saved the sum of things for pay". This concept brings to mind Albert Einstein's proposal in June 1947 to the US Secretary of State, George C. Marshall, for the establishment of a World Government. Einstein hoped for US leadership with bipartisan political concurrence. The UN, he said, had been ineffectual and could not control the situation. He expressed a sense of urgency, declaring that "the world is heading for an atomic war within 2 to 10 years" (*sic*) and that the destruction of civilization could be averted only by an effective supranational World Government to which military power would be transferred.

The 389 pages between the preface and the final chapter are a collection of essays generally related to the subject of the applied atomic era, which may be dated from about 1942 with the initiation of the Manhattan Project, and to the Cold War. Extended treatment is given to the 1952 and 1953 Nevada tests including Upshot Knoch, "the dirtiest and most hazardous detonations ever conducted in the United States" from which downwind fallout exposed sheep and shepherds to severe radiation. Special attention is devoted to the tragic experiences of Navajo uranium miners. As a lawyer, Udall represented them and their widows over many years in unsuccessful damage claims against the government, brought in the federal courts and ultimately rejected by the Supreme Court. He was devastated and sick at heart about the outcome, excoriating the courts and "the government in Washington that had betrayed" his clients.

Readers of Udall's fulminations on the hydrogen bomb and on the decision to use the atomic bomb against Hiroshima would do well to titrate his acerbity by consulting the contemporary and subsequent expressions of those who were involved in the discussions and decision-making within the government. A good start would be



*Men and Decisions* by Lewis L. Strauss (1962) both for what he says and especially for his references to specific contemporary documents and quotations from them. Strauss was one of the original members of the AEC, and later became its chairman. He was one of the overruled minority, including Under Secretary of the Navy Ralph Bard (and, to a degree, Secretary of War Stimson), who advocated warning the already defeated but not yet capitulated Japanese and then demonstrating the atomic bomb over an uninhabited area (specifically, for example, over a forest near the village of Nikko on the main Japanese island).

It is all very well in retrospect to see that the Nazis were nowhere near the United States in atomic-weapon development (although the Soviets proved to be dangerously close indeed), and to say decades later, after the collapse of the Soviet government, that the US "cold warriors' contempt for restraint had poisoned our politics," "altered American ethical standards and corrupted core institutions of our government" under the leadership of "two U.S. Presidents, imbued with military machismo". But could the presidents, the secretaries of state and war (and later of defence) and the atomic-energy commissioners responsibly have taken any risk of being outgunned by the Soviets?

Until their collapse, the Soviets appeared to be a clear and ever-present danger. True, this risk-avoidance led to escalation, because the Soviets feared the Americans as the Americans feared them. But this immensely costly escalation, including the Strategic Defense Initiative (Star Wars) may have partly determined the ultimate bankruptcy and collapse of the Soviet empire. Finally, the Soviets threw in their hand and dropped out of the competition. Has one erred and been profligate in paying fire insurance premiums if one's house does not burn down? Udall, with hindsight, ignores or dismisses such considerations.

What would the consequences have been if the Soviet Union had developed a hydrogen bomb and the United States did not have one? That was the underlying hazard confronting the decision-makers. It came into sharp and urgent focus in August 1949 when Soviet detonation of a fusion bomb was detected by the United States' recently created Long Range Detection System. Only four years later, the Soviets exploded a thermonuclear (hydrogen) device, less than a year after the successful first US test. Think of the leverage the Soviets would have had against the West if only they had had such an overwhelming weapon. Contrary to the impression given by Udall, a great debate flourished on this issue in the public media, as well as in classified and unclassified form in the government, through the

intervening months before President Truman gave the order on 31 January 1950 for the AEC to proceed with urgency to try to develop the 'Super', as it was called — the hydrogen weapon.

Sublimely self-assured and self-righteous, innocent of humility, Udall dismisses these practicalities and in his frequent animadversions attacks the motivations, the characters and the wisdom of the no less patriotic Americans who were on the stressful front lines of decision-making during the Second World War and the succeeding Cold War. *Caveat lector.* □

*William T. Golden, treasurer of the American Association for the Advancement of Science, is at 40 Wall Street, New York, New York 10005, USA. He served on the staff of the US Atomic Energy Commission for its first three years (as assistant to Commissioner Lewis L. Strauss).*

## String theory

Ronald Brown

### **The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots.**

By Colin C. Adams. W. H. Freeman: 1994. Pp. 306. \$32.45, £23.95.

How old are knots? It has been suggested that the Stone Age should be called the Age of String. The extraordinary tasselled figures photographed and described by G. L. Walsh in *Bradshaws: Ancient Rock Paintings of Western Australia* (Edition Limitée, 1994) have been suggested to be 50,000 years old. Knots have been intimately linked with the development of humans, through weapons, fishing, hunting, clothing, housing, boating and a myriad of other ways. The metaphor of knots is found throughout literature, and knots and interlacing are featured in many forms of art.

Some ancient problems, such as the slippage of knots, are still not analysed. Yet there is now an extensive mathematics of the forms of knots. It began with Lord Kelvin's romantic vortex theory of the atom. He had been fascinated by the properties of smoke rings shown by his friend Peter Guthrie Tait. His paper of 1867 states: "Diagrams and wire models were shown to the [Royal] Society [of Edinburgh] to illustrate knotted or knitted vortex atoms, the endless variety of which is infinitely more than sufficient to explain the varieties and allotropes of known simple bodies and their mutual affinities". The first step in this programme was to show how the variety of knots might model the periodic table. The analogy with atoms soon became lost, but Tait continued his work on classification of

knots, publishing three fundamental papers, and coming up with conjectures some of which, for example the flying conjecture on alternating knots, have only recently been settled, in this case by W. Menasco and M. B. Thistlewaite in 1990.

Knot theory developed with the rise of topology in the late 1920s and 1930s, and then continued to be attractive to a dedicated band for its combination of geometrical matter and algebraic method, although regarded by others as a specialist and marginal field. One of its main workers, H. Schubert, abandoned knot theory in the 1950s because of its apparent specialist nature, and took up category theory, seemingly the other extreme style of mathematics, with its discussion of broad theories and general methods. Unfortunately, he died before he could learn of the links that have now been made between the two subjects and methods. It is an indication of the growth of the area that this link is but one topic that Colin Adams's book cannot find space to treat.

A major breakthrough in 1928 was the polynomial invariant named after its discoverer, the US topologist J. W. Alexander. This allowed for the easy distinction of many knots and links of up to ten or eleven crossings, but failed to distinguish between the left- and right-handed trefoil. A second burst of activity in the 1980s was founded on work by the New Zealander Vaughan Jones. His polynomial invariant, which arose from his work in functional analysis, distinguished the left- and right-handed trefoils. It led to a slew of polynomials found by groups of workers all over the world. An important paper in 1985, which announced new polynomials, had five authors from three research groups, who had come to the same answer from different areas of mathematics, while a Polish team missed out through not being quick enough into print.

There are now connections with the uncoiling of DNA, which has a habit of getting knotted and uses specific enzymes to unknot itself for duplication. Knotted orbits are found in some chaotic motions. Research teams in Tokyo, and San Diego, California, have now found connections between knots and energy, both elastic and electrostatic. There are strong relationships with statistical mechanics.

*The Knot Book* sets out to demonstrate the excitement of mathematics, and to introduce scientists from a variety of fields to what is going on in knot theory. It is also usable as a textbook that will stimulate the imagination. It succeeds in these aims admirably, although for a mathematics course there are several topics that are omitted, such as the fundamental group and a full treatment of the classification of surfaces; for these the recent book by N. D. Gilbert and T. Porter, *Knots and Surfaces* (Oxford University Press, 1994),

would be more suitable.

An advantage of knot theory for explaining the methods of mathematics is that the problems can be stated simply, but solving them often requires a gamut of modern, but sometimes startlingly simple, ideas. Another feature is the use of analogy. For example, there is a way of composing knots and there is a notion of prime knots. So we have this extraordinary analogy between the way knots behave and the way numbers behave. But although there are an infinite number of prime knots, the proof of this is in no way analogous to the proof of the corresponding result for numbers.

Adams is an expert in knot theory, and this shows in the clarity and accuracy of his writing, and in the rich store of examples

and problems. A nice feature of the book is that, unusually for mathematics books at the undergraduate level, some of the topics are still the subject of current research, such as almost-alternating knots.

Many applications are given: there is a whole chapter entitled "Biology, Chemistry and Physics". A chapter on higher-dimensional knots stirs the imagination, and there are problems of varying levels of difficulty for the reader to enjoy tackling. We are going to see much more of knot theory and its applications, and this book is an excellent place to start. □

*Ronald Brown is in the School of Mathematics, University of Wales, Bangor, Gwynedd LL57 1UT, UK.*

## A geographer on well-trodden ground

*Peter Francis*

**Volcanoes.** By Alwyn Scarth. UCL Press/Texas A&M Univ. Press: 1994. Pp. 273. £40, \$45 (hbk); £14.95, \$19.95 (pbk).

THERE are many ways of knowing the world. In these politically correct times, scientists have learned that the perceptions of Earth held by other cultures may be as intrinsically valid as their own: science cannot be separated from the society in which it is carried out. Even within Western science, different disciplines have divergent perceptions of the same phenomena. Thus, whereas volcanologists passionately study volcanic ashes and lavas in order to seek an understanding of eruption mechanisms, geochemists view these merely as convenient samples from Earth's interior, and geophysicists ignore them entirely.

Most volcano books are written by authors with geological backgrounds. It is intriguing, therefore, to encounter one whose author explicitly states that he offers the perspective of a geographer. How might such a text differ from a conventional one? I expected an emphasis on the many ways in which volcanoes affect human society, not only directly through the obvious impact of eruptions, but also, for example, through their importance for agriculture and as sources of economic mineral deposits.

Alwyn Scarth has missed an opportunity here. Rather than opening a window on aspects of volcanoes generally skimmed by his geological counterparts, he has preferred instead to revisit much ground already well trodden by them. In common with most other texts, he offers descriptions of the various kinds of volcano and volcanic eruptions, and several narrative accounts of eruptions. Although he brings a freshness to some rather routine

material, he is ill-prepared to take on geologists on their own turf, and his book contains some egregious errors and curious placements of emphasis. He has much to say, for example, about the scoria cones built up by minor strombolian eruptions, but very little about the huge eruptions that form calderas tens of kilometres in diameter, such as Long Valley in California.

Any introductory level text that aims to avoid technical language, as Scarth's does, is likely to contain simplifications that may make subject specialists wince. These are forgivable if they lead to enhanced understanding. But Scarth's well-intentioned simplifications sometimes mislead. Readers might conclude, for example, that Pompeii was destroyed in AD 79 by the same kind of phenomenon that wrecked St Pierre (Martinique) in 1902, as Scarth

uses the term *nuée ardente* (glowing cloud) for both. Although their consequences were similar, the two phenomena had very different origins: St Pierre lay on the margins of a glowing avalanche of dense rock cascading down the flanks of Mount Pelée, whereas Pompeii was devastated by pumice-laden hurricanes that swept down from the plinian eruption column rising above Mount Vesuvius.

Although highly readable, Scarth's treatment of eruptions is often idiosyncratic. He had the opportunity to present the first account in a popular book of the massive 1991 eruption of Mount Pinatubo, surely one of the most important this century. But his treatment, although admirably clear, runs to only half-a-dozen pages and is inexplicably sandwiched between an account of the minor eruption of Heimaey (Iceland) in 1973 and a lengthy description of the 1730–36 eruption on Lanzarote, Canary Islands.

Such reservations apart, the book has much to offer. Scarth has delved deeply into the histories of his favourite volcanoes, particularly those of the Azores and the Auvergne, and his crisply written accounts are full of fascinating details. The story that he unearthed of the evacuation of Lanzarote by the Spanish in 1731 is a real gem. It is an instructive example of the dilemmas authorities have to resolve when confronted by an erupting volcano, a starving, frightened population and all the economic and political disruption that an evacuation engenders. Scarth would have produced a better, more distinctive and more useful book if he had used his background as a geographer to provide us with more new perspectives on human issues such as these. □

*Peter Francis is in the Department of Earth Sciences, Open University, Milton Keynes MK7 6AA, UK.*



Lava flow, one of over 100 photographs by D. Weisel that appear in *Fire on the Mountain: The Nature of Volcanoes*, with text by C. Johnson. Chronicle Books, \$29.95 (hbk), \$18.95 (pbk).



# Beefing about cattle farming

Verner Wheelock

**Beyond Beef: The Rise and Fall of the Cattle Culture.** By Jeremy Rifkin. *Thorsons (HarperCollins): 1994. Pp. 353. £8.99 (pbk). Published last year in the United States by Penguin.*

JEREMY Rifkin has collected a mass of information portraying all that is bad about beef production. According to him, there are today 1.28 billion cattle in the world. Cattle use land that could be much more efficiently used for other purposes. Much of the destruction of the Amazon

As a result of deregulation in the United States, federal meat inspectors have been stripped of most of their authority and examine fewer than 1 per cent of the carcasses. Because of this, they can no longer diagnose disease or identify contaminants.

Although Rifkin's emphasis is on the United States, there are some global dimensions. For example, the initial reason for developing beef production in the United States was to meet the demands of UK markets. Rifkin makes reference to an extensive list of sources. Some of these are publications by the US government, others are articles in newspapers and magazines.

As one might expect, Rifkin paints the blackest possible picture. Undoubtedly he strikes a chord, especially in the light of the outbreak of bovine spongiform encephalopathy in the United Kingdom; cases of infection with *Escherichia coli* 0157, a particularly virulent pathogen with beef one of its main sources; publication of the report of the Beef Tribunal in Ireland, which has confirmed the exist-

ence of many dubious practices within the industry; and the growing concern about animal welfare, especially during transport. A British visitor to Greece recently described in a radio interview how she had seen a lorry load of dead cattle that had suffocated because they were packed too closely together.

It would be wrong, of course, to deduce from this book that all cattle are badly treated and that all beef is a threat to human health. Nevertheless, Rifkin makes many valid points. The book will certainly provide ammunition for those who object to the use of meat in the human diet. It should be read by anyone who has an interest in food and in the way it is produced. Above all, it should be essential reading for all those involved in any aspect of the beef industry. If beef is to have a future, then the issues raised must be addressed and the allegations answered. □

Verner Wheelock is at Albert Mill, 10 Hey Street, Bradford, West Yorkshire BD7 1DQ, UK.

IMAGE  
UNAVAILABLE  
FOR COPYRIGHT  
REASONS

## Threat to human health?

rainforest during the 1970s has been attributed to large-scale cattle farming. The shortage of human food is exacerbated because of the large amounts of grain used to feed the cattle. About 60 million tonnes of methane are released into the atmosphere annually, thereby contributing to global warming.

The cattle themselves are subject to considerable ill-treatment, as the following allegation about US practice shows.

Because the journey to the slaughterhouse is often a rough and brutal one, animals frequently fall and are trampled upon inside the trucks, suffering broken legs or pelvis. Unable to rise these animals are known as 'downers'. Although downed animals are frequently in severe pain, they are rarely euthanised [*sic*] or anaesthetised, as that would translate into a lost carcass and additional expenses. Often spreadeagled on the floor of the trailers, unable to stand or walk, these hapless animals are chained by their broken legs and dragged from the truck onto the loading ramp to await their turn for slaughter.

# Strategies for living together

David L. Hawksworth

**Symbiotic Interactions.** By Angela E. Douglas. *Oxford University Press: 1994. Pp. 148. \$45, £29.50 (hbk); \$21.95, £12.50 (pbk).*

THE fundamental biological strategy of symbiosis is only rarely accorded prominence in considerations of ecology or ecosystem processes. Yet lichen symbioses are the major primary producers on eight per cent of the land surface, and algae in coral fix three to four times more carbon than the phytoplankton in all nutrient-upwelling regions of the oceans. In adopting a 'process' rather than an 'association' approach, this book is immediately set apart from *The Biology of Symbiosis*, which the author wrote with Sir David Smith in 1987, and from other texts. Disparate organisms appear side by side in consideration of novel structures and metabolic capabilities, nutritional interactions and the formation and regulation of symbiosis. In a period of biology characterized by increasing specialization, it is refreshing to be able to read in the space of a few lines of luminous bacteria, fish and squids, herbivorous mammal gut processes, root nodules of legumes and lichen thalli.

But what should be the compass of such a title? Douglas prudently rejects mutual benefit as a criterion for symbiosis, and defines symbiosis as "associations between different species that persist for long periods (relative to the generation

time of the interacting organisms)". Is this really what de Bary intended in 1879 in explicitly including parasitic relationships?

And what should the partners be called? Douglas refers to the larger organism as the 'host' and the smaller as the 'symbiont'. Contrary to the treatment in *The Biology of Symbiosis*, this terminology is considered inappropriate for mycorrhizal fungi and plants, but should it really be maintained outside host-parasite relationships? In the case of lichen thalli, the fungal partner is treated as the 'host', yet the structure results only from the interaction, and de Bary himself regarded the algae in lichens as hosts to the fungus. Adoption of the neutral terms 'inhabitant' and 'exhabitant' proposed by R. Law and D. H. Lewis in 1983 would have avoided any need for assumptions about relationships, and enabled each partner to be referred to as a 'symbiont'.

This book is most timely. Symbiotic associations are particularly at risk from the loss of biodiversity, as in many cases one inhabitant can form associations with a variety of exhabitants. The complexity, and our level of ignorance, of such associations spring from this book. It will stimulate both aspiring and established biologists. □

David L. Hawksworth is in the International Mycological Institute, Bakeham Lane, Egham, Surrey TW20 9TY, UK.

# Microstructured semiconductor lasers for high-speed information processing

P. L. Gourley

**The technology to generate laser light from semiconductor microstructures is developing rapidly. New techniques for epitaxial growth and surface processing allow great flexibility in laser architecture. Highly efficient lasers with submicrometre dimensions, operating independently or packed together in dense two-dimensional arrays, will soon be used for transmitting, storing and manipulating information with unprecedented speed.**

IN C. S. Lewis's *The Magician's Nephew*, the world of Narnia began with the creation of lights by the great voice: "the blackness overhead, all at once, was blazing with stars. They didn't come out gently one by one, as they do on a summer evening. One moment there had been nothing but darkness; next moment a thousand, thousand points of light leaped out."

Like the creation of the vast number of stars in Lewis's novel, modern photonics engineers are creating millions of microscopic laser lights on gallium arsenide semiconductor chips. These chips comprise periodic arrays of micrometre-sized lasers emitting coherent visible or near infrared light. These lasers can operate as large arrays or independently to communicate millions of messages at the same time. Such vast numbers of light-emitters could be used to read, write or process two-dimensional images, and to speed the flow of information between memory and processing chips. By replacing cumbersome electronic wires with laser beams, bits of data could move faster between spots on different chips.

Developing semiconductor laser materials use alloys like AlGaAs or other compound semiconductors which emit wavelengths of light extending from 450 nanometres in the blue to 10 micrometres in the infrared. This range covers the wavelengths of the most familiar 'benchmark' lasers such as argon ion (514 nm), YAG (1.06  $\mu\text{m}$ ) and CO<sub>2</sub> (10.6  $\mu\text{m}$ ). Each of these familiar industrial lasers can produce hundreds of watts of continuous light power. Arrays of the most developed semiconductor lasers, emitting 870-nm light from cleaved GaAs wafers, have demonstrated similar power levels. Discrete semiconductor lasers typically emit from 0.1 mW to 1 W and are the size of a salt grain, many orders of magnitude smaller than benchmark lasers. Semiconductor lasers are clearly the most efficient light sources, with up to 50% overall quantum efficiency, far exceeding efficiencies of the familiar industrial lasers (0.1–0.01%), tungsten lamps (few per cent) and fluorescent arc amps (10%).

Beyond highly efficient light generation, semiconductor lasers are well suited to high-volume manufacturing. Like the integration of silicon transistors in the 1960s, the semiconductor laser will undergo integration in large arrays and with other optical circuit elements such as waveguides and detectors. The ultimate uses of optical circuits are still an open question. But it is likely that information display and transfer, energy delivery and computation are among the leading applications. As these applications become a reality, our society will be more dependent on new materials for photonics.

## Three-dimensional architectures

Many technological advances are inherited from the new geometries of microstructured semiconductor lasers<sup>1</sup>. These new forms have emerged through modern semiconductor microfabrication, including epitaxial growth and surface processing. Such technologies allow the creation of three-dimensional laser architectures that enhance light emission from the materials used to build them.

Several advantages follow from the microstructure inherent in the crystalline growth of the laser crystal. The crystal is literally built up atomic layer by atomic layer onto a semiconductor substrate by epitaxy. Techniques like molecular beam epitaxy and metal organic vapour-phase epitaxy allow sequential layering of hundreds of different semiconductor alloys. By mixing and matching these 'designer materials', crystals can be layered together to optimize the electrical and optical properties. This tailoring of semiconductor structures is called bandgap engineering, and is described below.

Every semiconductor has an energy gap that separates lower-energy, occupied electron states (valence bands) from the higher-energy, empty states (conduction bands). An electron promoted across this gap will result in a mobile 'hole' in the valence band and a mobile electron in the conduction band. In bandgap engineering, different semiconductor alloys with different bandgaps are joined during growth to form a single-crystal heterojunction. This heterojunction creates potential-energy steps in conduction and valence bands that inhibit motion of electrons and holes. A double heterojunction with a thin layer (~100 Å) of low-bandgap material (such as GaAs) sandwiched between two wide-bandgap materials (such as AlGaAs) is called a quantum well (Fig. 1 shows a three-quantum-well structure); in these structures, used in most semiconductor lasers, individual electrons and holes are confined within the central layer.

An alternative, thicker heterostructure is an optical waveguide layer, in which a thick (~1,000-Å) GaAs layer is sandwiched between two AlGaAs layers. Light is guided by total internal reflection along the GaAs layer because it has a higher refractive index than AlGaAs.

Another important heterostructure is a quarter-wave mirror stack, sometimes called a distributed Bragg reflector. It comprises periodic pairs of layers (GaAs and AlAs, for example) with high and low refractive index to reflect light incident perpendicular to the layer (Fig. 1). The waves of light reflected from each interface add constructively to create a very high reflectance exceeding 99%. The individual layer thicknesses  $d$  correspond to a quarter wavelength of light,  $d = \lambda/4n$  where  $\lambda$  is the wavelength and  $n$  the refractive index. For GaAs  $n = 3.6$  at the emission wavelength of  $\lambda = 850$  nm, so  $d = 630$  Å.

With epitaxial growth, these building blocks (quantum wells, optical waveguides and mirrors) can be stacked in a single crystal according to a designer's wishes. After growth, the surface of the crystal wafer can be patterned by optical or electron-beam microlithography using a spin-on resist material. After the resist is exposed and developed, selected areas as small as a few tens of nanometres can be removed from the wafer's surface by etching techniques. These techniques use 'wet' chemicals or 'dry' ion beams to strip off layers with nearly the same precision as they were laid down.

Taken together, epitaxial layering and surface processing allow the photonics engineer great flexibility in the geometry of



lasers. Developing semiconductor lasers exhibit an astonishing variety of geometrical forms including multilayer sandwiches<sup>2,3</sup>, posts<sup>4</sup>, gridirons<sup>5</sup>, thumbtacks<sup>6</sup>, swiss cheese<sup>7</sup>, concentric rings<sup>8</sup> and honeycombs<sup>9</sup>. A checkerboard geometry<sup>10</sup> for a phase-compensated two-dimensional surface-emitting laser array is shown in Fig. 2a. The pattern of light emitted from an uncompensated array is shown in Fig. 2b and c, for the near and far fields respectively.

### Vertical-cavity surface-emitting lasers

Although many microstructured lasers are still in their infancy, others are more developed. The vertical-cavity surface-emitting laser (VCSEL) (Fig. 1) is now making the transition from research to manufacturing. This structure includes all the necessary laser components: optical cavity, gain medium and electrical pathway for current, built within a single-crystalline architecture. It comprises three basic components: an n-type mirror for conducting electrons, a p-type mirror for conducting positively charged holes, and a central quantum-well active region for optical gain. The mirrors define an optical cavity that confines photons for the lasing process.

The general operation of a semiconductor laser can be explained with Fig. 1. With a low voltage (a few volts) applied, electrons flow from the leftmost metal contact through the substrate and n-type mirror to quantum wells, where they collect. At the same time holes are injected through the rightmost metal contact, through the p-type mirror, and into the thin quantum wells. When an electron encounters a hole in the quantum well, the pair is annihilated (or recombines) and a photon is produced. The light produced by this chance encounter is called spontaneous emission. Spontaneous photons are randomly emitted in all directions and are frequently reabsorbed by the quantum well materials. Spontaneous emission tends to be weak and have a broad energy distribution.

When the bias voltage is increased, dramatic changes take place. Increased flow of electrons and holes fills the energy states of the quantum wells and optical gain is developed. In this condition, a spontaneous photon traversing the quantum wells will induce another electron-hole recombination event and photon emission (stimulated emission). These stimulated photons multiply the photon population many times. Only those waves which reinforce themselves after a cavity roundtrip with two mirror bounces will build up. These stimulated photons add constructively to produce an intense standing wave in the cavity. In the steady state, some of the stimulated photons leak out through the mirrors, producing a coherent, monochromatic beam outside the cavity.

Bandgap engineering of the vertical-cavity laser structure has evolved over the last fifteen years<sup>2,3,11-23</sup>, following the pioneering work by Iga of the Tokyo Institute of Technology<sup>2</sup>. Most of the pioneering VCSEL diode work has been accomplished with the GaAs/AlGaAs materials that emit in the near infrared (770–1,000 nm). Recent advances include demonstration of AlGaInP-based visible (red at 650–670 nm) VCSELs (Fig. 3)<sup>21</sup>, and InGaAsP-based VCSELs operating at 1.3  $\mu\text{m}$ . Newly developed compound nitride semiconductors may even allow short-wavelength VCSEL lasers emitting in the ultraviolet-blue region of the spectrum.

### Advantages of microstructured lasers

The vertical cavity of Fig. 1 is versatile; it can be used for light generation<sup>2,3</sup>, modulation<sup>24</sup> or switching<sup>25,26</sup>. As a laser, it can

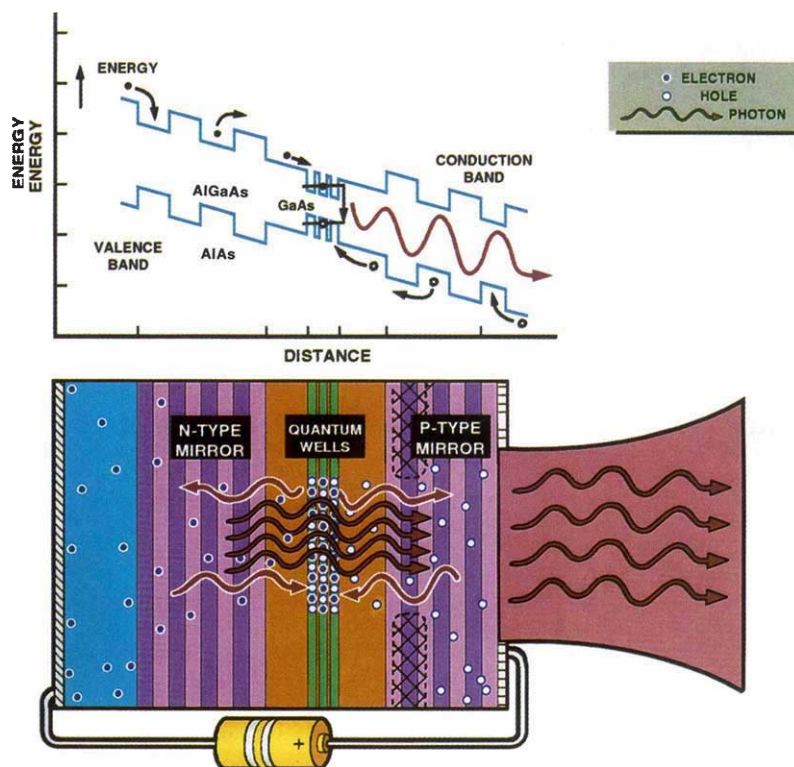


FIG. 1 Diagram of a single vertical-cavity surface-emitting laser. The upper part shows the conduction and valence bands in which the electrons and holes move. The bottom part shows the layered pairs of semiconductor materials used for the mirror region and thinner quantum-well layers used in the active region. Electrons and holes are shown by dark and open circles respectively. These carriers are injected from metal contacts applied to the substrate and top mirror layer. The carriers collect into the quantum wells where they recombine to produce photons (wavy arrows). Photons with the right wavelength are multiplied inside the laser cavity. Some of these photons leak out the top mirror, through an aperture defined in the top metal contact, to produce an intense beam.

operate as a planar device, etched post or two-dimensional array. Its architecture is radically changed from the conventional edge-emitting laser invented in the early 'sixties. In contrast to that laser, which emits widely diverging light from a cleaved wafer-edge, the vertical-cavity laser emits light normal to the wafer surface. This allows vast numbers of lasers to be packed side-by-side on the wafer surface or to be integrated vertically with other optical elements. It also allows the laser aperture to shape the emerging beam into a low diverging circle, compatible with fibre optics. The circular beam cross-section of a VCSEL is contrasted with the stripe-like beam from an edge-emitter in Fig. 3.

The advantages of bandgap engineering are illustrated by a recent scheme to lower the electrical resistivity of VCSEL mirrors. The quarter-wave stack must serve three functions: as a highly reflecting mirror, as a conducting path for charge carriers and as a sink to remove unwanted heat. These three requirements are not generally compatible, but can be accommodated by bandgap engineering. For the first function, large differences in the indices of refraction of adjacent mirror layers give high reflectivity. Large index differences are accompanied by large energy barriers, which charge carriers must surmount as they move across layers. These barriers impede carriers, causing high resistances which create heating and lower efficiency. To lower the resistance but maintain high reflectivity, the interfaces between layers are graded in alloy composition to 'soften' the energy barrier. A grading approach developed at Sandia<sup>27</sup> and at the University of California at Santa Barbara<sup>28</sup> have produced low-resistance mirrors enabling VCSEL threshold voltages of less than 1.5 V and efficiencies of up to 30%.

The laser benefits from bandgap engineering of the active region. The quantum wells can be placed commensurately with

FIG. 2 *a*, Scanning electron micrograph of a two-dimensional vertical-cavity surface-emitting laser array with integrated phase compensator made at Sandia National Laboratories by Warren and co-workers. The individual laser pixels are 5- $\mu\text{m}$  squares with every other pixel covered with a glass-shifting layer. *b*, The near field laser light emitted from an uncompensated  $20 \times 20$  array injected with current. *c*, The four-lobed light pattern observed far from the device surface (far field).

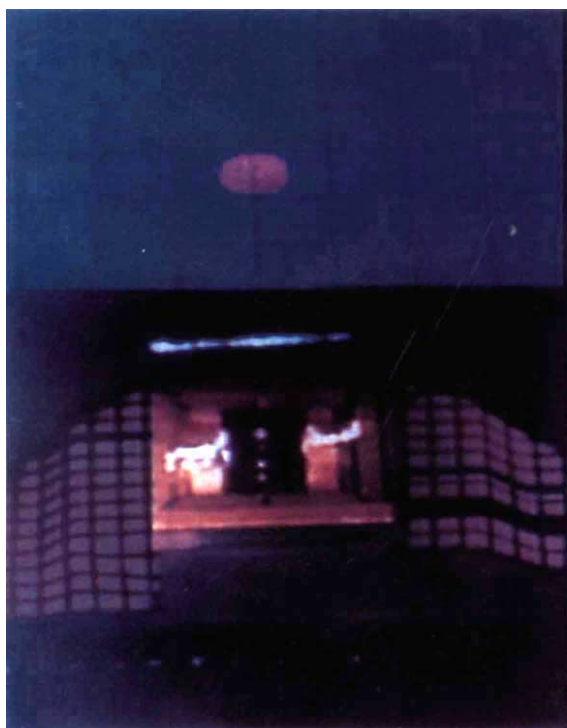
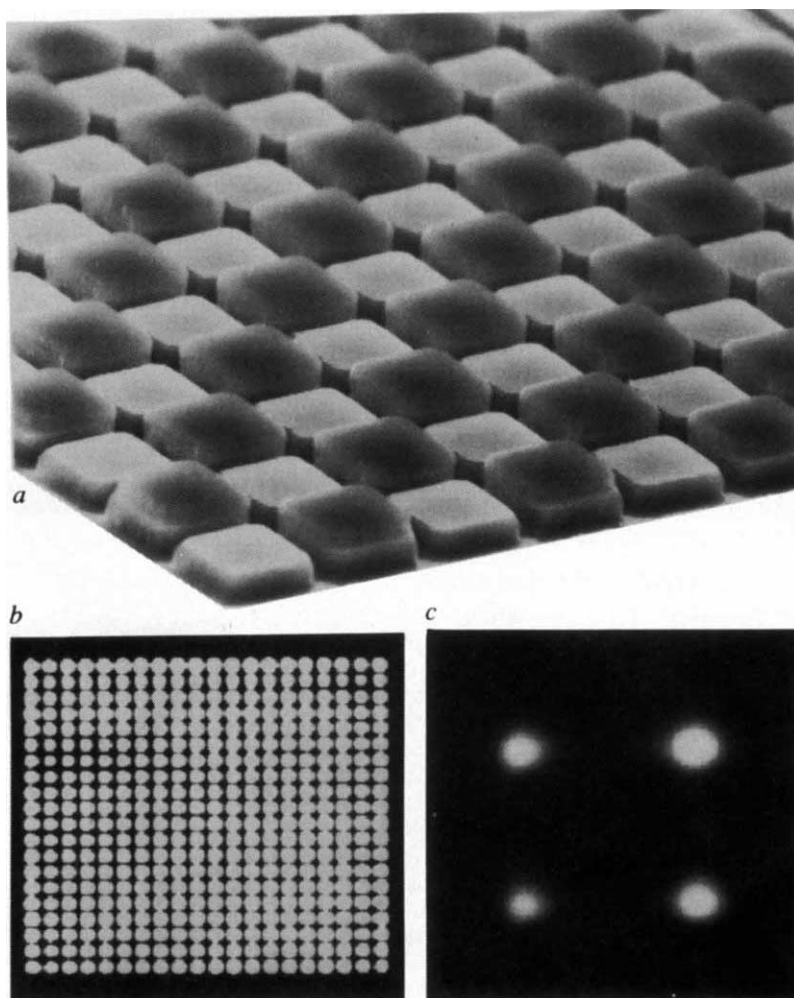


FIG. 3 Cross-sections of visible laser beams from a new surface-emitting laser diode (left) contrasted with conventional edge-emitting laser diode (right). The red beam from the surface-emitter illuminating the backside of a paper screen is

circular and much smaller than the wide, stripe-like beam from the edge-emitting laser. These properties simplify the collimation and focusing of the VCSEL. (Photographs courtesy of R. Schneider and K. Lear of Sandia National Laboratories.)



the antinodes in the optical standing wave<sup>29-31</sup>. This effect is illustrated in Fig. 1 where the crest in the optical waves aligns with the quantum wells. The lasing wavelength is stabilized, the gain material is more effectively utilized to feed the lasing process, and evidence suggests that spontaneous emission is enhanced<sup>32,33</sup>. The quantum wells can be thin elastically strained layers, to extend lasing wavelength and increase efficiency<sup>34-38</sup>. Furthermore, the ultra-short active region can be designed to improve temperature characteristics<sup>39,40</sup> and increase the upper frequency ( $\sim 14$  GHz) for switching the laser on and off<sup>41-45</sup>.

Beyond performance characteristics, the vertical-cavity surface-emitting-laser is particularly attractive for manufacturing. VCSELs can be individually tested while they are still in wafer form. Because defective lasers are identified before packaging, costs are reduced. The surface emission geometry of VCSELs will also simplify packaging.

## Surface-patterned laser arrays

Whereas layered microstructure is used to enhance the efficiency of the lasing process, surface microstructure is used to shape and direct the beam emitted from the wafer. Surface patterning is useful for creating two-dimensional arrays of surface-emitting lasers. When many of these tiny lasers are placed side by side on a wafer, it is possible to operate them in a phase-locked condition (Fig. 2b). As a consequence, the entire array operates in unison to create a narrow laser beam. By adjusting the relative phase of the lasers, it may even be possible to steer the beam in space at very high speed.

The lasing modes in a two-dimensional arrays have been studied recently<sup>5,10,46-51</sup>. The optical mode in the array will determine the angular distribution of light observed far from the array surface (far field). The far-field patterns observed experimentally reveal four bright spots (Fig. 2c). The spots are placed symmetrically  $5^\circ$  away from the optical axis into each of the four quadrants of the far field. The width of the spots ( $\theta$ ) is about  $2^\circ$ , slightly broader than predictions of  $\theta \approx \lambda/D$  by diffraction theory, where  $D$  is the diameter of the array. The surface-patterned array has a beam with much lower divergence than that of a single element of the array.

Despite the advantages of narrower divergence, the beam intensity is off the optical axis (normal to wafer surface). Most applications require a single beam on the optic axis. The off-axis beam is due to the array mode which corresponds to each element being  $180^\circ$  out of phase with its neighbours. This antisymmetric mode dominates the lasing process. The mode amplitude has a zero crossing between elements, consistent with the quenching of lasing in the channels. Attempts have been made to produce a symmetric mode with on-axis intensity, but with little success.

A more successful approach taken by Warren and co-workers is to let the array lase in its preferred antisymmetric mode and then use epitaxial binary optics to adjust the phase in the near field<sup>10,49</sup>. In this case, an extra phase-correction layer is grown above the top mirror. After growth, the layer is selectively removed above every other lasing element. This leaves a checker-board pattern etched into the phase-correction layer (Fig. 2). In the squares with remaining material, the phase of the outgoing wave is retarded by  $180^\circ$  with respect to the squares without the material. The phase front of the wave is now compensated across the entire array surface, as if the array was operating in a symmetric mode. The far-field pattern comprises a bright lobe on

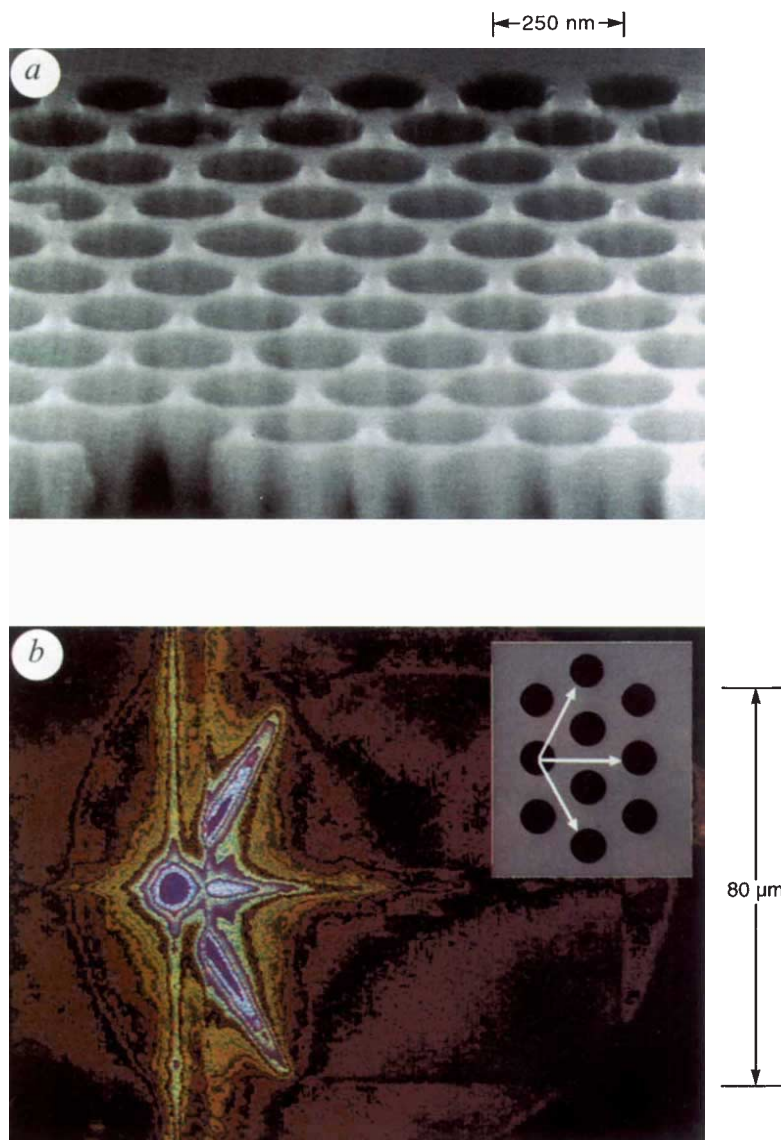


FIG. 4 a, Scanning electron micrograph of a photonic lattice structure fabricated with electron beam lithography and reactive ion beam etching by Wendt and Vawter at Sandia National Labs. The structure features 250-nm diameter holes etched into a AlGaAs multilayer wafer. Photonic lattices such as this one are being investigated for future generations of low-threshold lasers. b, A light scattering micrograph showing how light propagating along the wafer plane is redirected out of the plane by the lattice.

the optic axis as desired. There are also secondary lobes off the optic axis, but these are weaker.

## Photonic lattices and band gaps

The above example shows how surface patterning of an epitaxial layer can be used to produce a phase transformation in a wave exiting the crystal. A slight phase adjustment in the near field produces a dramatic change in the far field, which is its Fourier transform. Study of the far field can provide valuable information about the optical modes in the arrays<sup>50</sup>. These arrays form a subset of more generalized periodic optical structures called photonic lattices. A photonic lattice is formed by periodic modulation of the dielectric constant (or refractive index) in one, two or three dimensions. The distributed Bragg reflector is a one-dimensional photonic lattice. Lattices of all dimensionalities serve to inhibit the propagation of electromagnetic radiation over a range of frequencies called a stop band. A stop band in three dimensions could entirely prohibit spontaneous emission. These effects would be useful for laser cavities, waveguides and detectors.

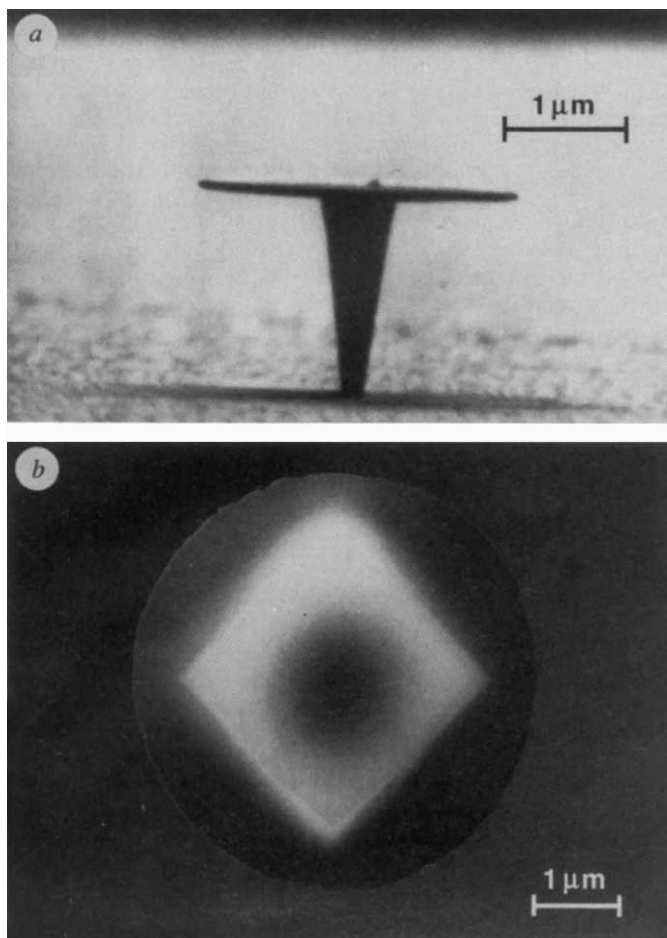


FIG. 5 Scanning electron micrographs of microdisks in the side view (a) and top view (b). The InP pedestal is a rhombus in cross-section as seen in the top view. It tapers to smaller dimensions as it approaches the substrate. (Photograph courtesy of R. Slusher of AT&T Bell Laboratories.)

Optical waves in the lattice are analogous to electron Bloch wavefunctions in a crystalline solid<sup>52-55</sup>. The electron wavefunction in the solid is described by local orbital motion around each atom and by long-range wave motion throughout the periodic arrangement of atoms. When the electron wavelength is twice the atomic spacing, the Bragg condition for reflection occurs. Under this condition, the electron wave cannot propagate,

creating the energy bandgap described earlier in this article. In the photonic lattice the stop band corresponds to the energy bandgap. Like electrons, the photons can exist in energy states where propagation is allowed. Unlike electron states which accommodate one electron at a time, the photon states can accommodate many photons at a time. These differences endow the photonic lattice with unusual properties, useful for nonlinear optics and not found in electronic lattices.

Recently, specific three-dimensional photonic crystals have been proposed by Yablonovitch<sup>56</sup> and others<sup>57-60</sup>. These structures are being intensely studied theoretically by many groups. This work has shown that complete three-dimensional photonic bandgaps are possible for certain crystal symmetries with sufficient dielectric modulation. Experimentally, the situation is much less developed. Experiments performed at microwave frequencies have confirmed the existence of photonic bandgaps in two dimensions (posts in triangular arrays<sup>61</sup>) and three dimensions (stacked-log structures operating up to 500 GHz<sup>62,63</sup>) using insulators or semiconductors. Structures with bandgaps in the optical frequencies have proven difficult to make because of the small submicrometre dimensions necessary.

It may be easier to make two-dimensional structures with photonic bandgaps at optical frequencies. Two-dimensional lattices which have a complete in-plane photonic bandgap have been theoretically predicted by Meade *et al.*<sup>64</sup> and Villeneuve and Piche<sup>65</sup>. Honeycomb lattices and triangular lattices of air holes have a complete gap for both transverse electric and transverse magnetic polarizations. The maximum bandgap occurs in structures with ~85% air-volume fraction. Honeycomb nanostructures have been fabricated by Wendt and Vawter at Sandia National Laboratories<sup>9</sup>. These structures, (Fig. 4a), were fabricated in compound semiconductors using electron-beam lithography and reactive-ion-beam etching. The honeycomb nanostructure corresponds to two-dimensional second-order Bragg-reflector wavelengths near 830 nm. This includes wavelengths for the spontaneous emission from GaAs quantum wells. Initial optical characterization of these lattices reveals that light normal to the plane can be resonantly coupled into the lattice near the Bragg condition<sup>9</sup>. A photomicrograph of light propagated and scattered out of the lattice is shown in Fig. 3b. The image shows that the lattice is effective in controlling the direction of photon propagation.

### Whispering-gallery microdisk lasers

One major hurdle for semiconductor lasers is the elimination of the heat that accompanies light generation. To obtain high operating efficiencies, considerable energy must be expended to overcome the threshold condition for inverting the electron populations in order to start the lasing process. This threshold

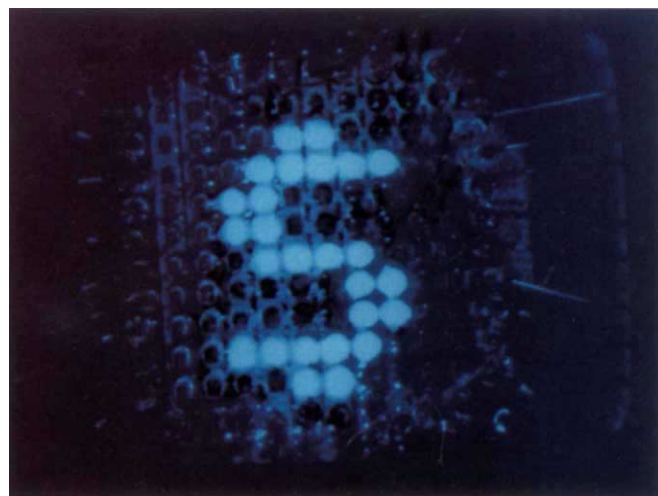


FIG. 6 A 850-nm light display from a  $10 \times 10$  matrix addressable array fabricated at the Solid State Technology Center of AT&T Bell Laboratories. Matrix addressing has the advantage of minimizing the number of wires to inject current. The dollar-sign display is created by raster scanning the array to turn individual elements on and off. These arrays are proposed for applications such as video imaging, confocal microscopy and spatial light modulation. (Photograph courtesy of R. Morgan, formerly with AT&T Bell Laboratories and now with Honeywell.)



condition might be compared to the heat required to raise the temperature of a pot of water to produce boiling. If the pot were made small enough and the energy supplied without loss, the threshold energy for boiling could be very low. How to make such a 'thresholdless laser', suggested by Yamamoto<sup>66</sup>, is a topic receiving much attention. Such a device would operate at the quantum limit: the first injected electron-hole pair would produce a laser-like photon. One approach to implementing this laser uses a photonic lattice<sup>9</sup> (Fig. 4a).

Other approaches to confine light for thresholdless lasers are also being investigated<sup>7</sup>. These approaches include fabrication of vertical-cavity posts, hemispherical cavities and whispering-gallery disks<sup>6</sup>. The object is to make small high-quality-factor microcavities to localize a few modes from the surrounding free-space modes. If the cavity is small enough, it can even modify the spontaneous emission rate of photons. Useful spontaneous emission can be enhanced and competing emission can be suppressed. The result is that device quantum efficiency increases. Like vertical cavities, these structures feature single-mode lasing, high-wavelength stability and low temperature-sensitivity. Ultrasmall cavities may even produce photon number squeezing, where the emitted photon stream is highly correlated in time and does not obey Poisson statistics. Theorists are now thinking of schemes to produce a correlated flow of electrons to inject such a microcavity.

A measure of merit for these microcavities is the fraction of spontaneous emission that is coupled into the lasing mode ( $\beta$ ).  $\beta$  can also be interpreted as the inverse of the number of optical modes in the cavity. As  $\beta$  approaches 1, the threshold current for lasing decreases, the lasing linewidth narrows, and the laser intensity fluctuations decrease<sup>66</sup>. An approximation for the spontaneous emission coefficient at wavelength  $\lambda$  is  $\beta \approx (\lambda/2n)^3 \lambda / V \Delta\lambda_{sp}$  where  $(\lambda/2n)^3$  is the optical volume per mode,  $V$  is the total cavity volume and  $\Delta\lambda_{sp}$  is the linewidth of the spontaneous-emission spectrum. Large-area vertical-cavity surface-emitting lasers have  $\beta \approx 0.01$ . Posts etched into these laser wafers have  $\beta \approx 0.1$ . Even higher values ( $\beta \approx 0.3$ ) are attained with 'whispering-gallery' microdisk lasers<sup>6</sup>. The values in photonic lattices and photonic bandgap materials are expected to approach 1, but have not yet been demonstrated experimentally.

A scanning electron micrograph of a 'whispering-gallery' microdisk laser is shown in Fig. 5. These microdisks were fabricated by etching quantum-well layers of InGaAs/InGaAsP (100 Å well/100 Å barrier) grown on InP substrates. The total thickness is 500 Å for a single-quantum-well structure and 1,500 Å for a six-quantum-well structure. Photolithography was used to pattern cylinders with diameters of 3–10 µm. A hydrochloric acid solution was used to etch away the InP around and below the protected circles, leaving a very thin quantum well/barrier disk. The disk is mechanically stable and lases when optically or electrically pumped.

These microdisk lasers are simple to make and attractive for several reasons. First, the disk volume is extremely small, only 1 µm<sup>3</sup>. This is about 100 times smaller than a vertical-cavity laser. Consequently, the current required for lasing is very low, in the microamp range. Second, the refractive-index mismatch between the semiconductor disk ( $n=3.6$ ) and surrounding air ( $n=1$ ) is enormous. This leads to very high confinement of spontaneous or stimulated photons created inside the disk. Only a single transverse optical mode is trapped in the thin disk. Third, the photons tend to skim along the circumference of the disk where they experience total internal reflection as in an optical waveguide. This gives very high overlap between the optical wave and gain present in the quantum wells, allowing the laser to operate more efficiently.

The optical waves in these disks are referred to as 'whispering-gallery' modes, analogous to sound waves first observed by Lord Rayleigh. More than a century ago, he explained how conversations could be heard at opposite walls inside the great dome of St Paul's Cathedral in London. The conversation was channelled

to the listener by efficient propagation of sound waves skimming along the walls. Similar optical waves have been observed in polymer spheres and water droplets. Light can be highly confined within these microstructures because of the high index mismatch between the sphere and its surroundings.

Although light is tightly confined in microdisk lasers, sufficient leakage occurs through the sides to form an output beam. The output power ranges from tens to hundreds of microwatts. The light emitted from a 5-µm microdisk is spread into a  $\sim 20^\circ$  angle about the circumference of the disk. Exactly how to best utilize this circular spread of light is a topic of high interest. The most useful output may emit from an array of microdisks, or from microdisks that are specially positioned to reflect light from the substrate.

As microdisks and other laser structures shrink to smaller and smaller dimensions, physicists are becoming especially intrigued by fundamental limitations on laser operation. The electron, photon and phonon (lattice vibrations) states change from continuous distributions to discrete levels separated by large energies. Electrons and holes injected into such structures find fewer degrees of freedom to relax excess energy. As a result, the energy distribution of carriers is 'hot' and spread over a large range of energies. This energy spreading lowers the conversion efficiency of an injected electron-hole pair into a lasing photon. These limitations will present new challenges for scientists and engineers to overcome.

## Quantum-cascade lasers

The emission wavelength of most semiconductor lasers is fixed by the energy bandgap of the quantum-well material in the active region. The laser wavelength is  $\lambda = hc/E_g$  where  $h$  is Planck's constant,  $c$  is the speed of light in a vacuum and  $E_g$  is the energy gap between electrons in the conduction band and holes in the valence band. Recently, a fundamentally new microstructured semiconductor laser<sup>67,68</sup> has been demonstrated by Capasso and co-workers at AT&T Bell Laboratories which removes this constraint. This new laser, dubbed the quantum-cascade laser, has its wavelength set, not by the bandgap, but by smaller energy separation of conduction sub-bands in adjacent quantum wells of different thickness. Since the sub-band energies change with quantum-well thickness, the lasing wavelength can be changed merely by growing thicker or thinner quantum-well layers. Using 8-Å and 35-Å InGaAs well-layers sandwiched between 35-Å AlInAs barrier layers, the demonstration laser emitted light at 4.2 µm in the infrared. Lasers operating at wavelengths of 4–100 µm, appropriate for environmental monitoring and analytical spectroscopy, should be possible with different layer thicknesses.

This remarkable new laser relies on bandgap engineering to accomplish this feat. It was produced by molecular beam epitaxy of 500 layers of variously doped semiconductor alloy layers to define the quantum-well active region, waveguide cladding layers and top contact layer. The laser light travels along the layers in a waveguide and out-edge mirror-facets formed by cleaving the wafer. This lasing direction is similar to conventional edge-emitting laser diode light.

In contrast to both edge- and surface-emitters, this new laser is not a diode. In the diode lasers both electrons and holes are required for recombination events to produce photons. In the quantum-cascade laser only electrons in the conduction band are required. The electrons are injected over a barrier layer where they pile up in the narrow 8-Å InGaAs layer. The electrons cascade downward in energy to the 35-Å InGaAs layer by emitting a 4.2-µm photon. After releasing this energy, they are quickly swept away into adjacent layers. The rates of injection, photon emission and sweepout are carefully regulated by design of the bandgap-engineered layers. This careful regulation is necessary to maintain a population inversion of electrons between the energy levels of the narrow and wide InGaAs quantum-well layers, so the stimulated emission process can take place.

With stimulated emission, up to 8 mW output power at cryogenic temperatures was reported. It is likely that higher output power and operating temperature will be achieved in succeeding laser designs. The implications of mid-infrared to submillimetre-wavelength lasers based on wide-bandgap, technologically mature compound-semiconductor alloys are apparent. Low-bandgap materials for longer wavelength edge-emitting diodes are difficult to grow, hard to process and highly temperature-sensitive. The quantum-cascade laser materials are compatible with those used for long-wavelength infrared detectors based on optical absorption between quantum-well sub-bands. No p-n junction is required, so difficulties associated with p-type doping and hole transport are eliminated. Further, the temperature dependence of the threshold current is predicted to be much weaker than that in diode lasers.

## Applications of microstructured lasers

As researchers continue to develop new microlasers with smaller volumes and higher  $\beta$ , manufacturers are scaling up to produce the more developed VCSEL lasers for several different applications. American manufacturers include Motorola, Photonics Research Incorporated and Optical Concepts. One application of interest, a matrix addressable array display<sup>69</sup> developed at AT&T Bell Laboratories, is shown in Fig. 6. For fibre-optic communications, these lasers feature high coupling efficiency to fibres, low bit error rate, low temporal dispersion and a threshold current that is insensitive to changes in temperature. Such VCSEL applications are under investigation by the ARPA sponsored Optoelectronics Technology Consortium (OETC)

which includes AT&T, GE, Honeywell and IBM. The OETC is working towards a 32-channel, 500-Mb s<sup>-1</sup> optical data link for board-to-board communications<sup>69</sup>. The red VCSEL is also an excellent candidate for use in printing, read/write and scanning applications such as a replacement for He-Ne lasers in super-market bar code scanners. Reprographic films are more sensitive at red wavelengths, and the circular beam simplifies focusing light on the film. Plastic micro-optics can be integrated with the lasers, resulting in low-cost miniature laser systems. Finally, two-dimensional VCSEL arrays are particularly attractive for a wide variety of image processing and optical computing applications.

An unprecedented capacity to synthesize microstructures is changing the form of the semiconductor laser. The ability to confine single optical modes in submicrometre-sized lasers in a single crystal is also changing the way we think about semiconductor lasers. It is possible to envisage extremely high-confinement resonators occupied by single photons. Low quantum occupation would allow entirely new kinds of photonic devices using minute power. These devices could become viable system components for transmitting, storing and manipulating information. The conception, fabrication and study of these device structures is creating new opportunities and exciting new challenges for scientists and engineers. Vertical-cavity surface-emitting lasers have already moved toward the manufacturing stage. They are compact, efficient and relatively inexpensive to manufacture in large quantities. The future of these tiny lights appears to be very bright indeed. □

P. L. Gourley is at Sandia National Laboratories, Albuquerque, New Mexico 87185, USA.

1. Yamamoto, Y. & Slusher, R. E. *Physics Today* **46**, 66–73 (1993).
2. Iga, K., Koyama, F. & Kinoshita, S. J. *Quantum Electron.* **242**, 1845–1855 (1988).
3. Gourley, P. L. & Drummond, T. J. *Appl. Phys. Lett.* **50**, 1225–1227 (1987).
4. Jewell, J. L., Scherer, A., McCall, S. L., Gossard, A. C. & English, J. H. *Appl. Phys. Lett.* **51**, 94–96 (1987).
5. Gourley, P. L. et al. *Appl. Phys. Lett.* **58**, 890–892 (1991).
6. McCall, S. L., Levi, A. F. J., Slusher, R. E., Pearson, S. J. & Logan, R. A. *Appl. Phys. Lett.* **60**, 289–291 (1992).
7. Yablonovitch, E., Gmitter, T. J. & Leung, K. M. *Phys. Rev. Lett.* **67**, 2295–2298 (1991).
8. Erdogan, T., King, O., Wicks, G. W., Hall, D. G., Dennis, C. L. & Rooks, M. J. *Appl. Phys. Lett.* **60**, 1773–1775 (1992).
9. Gourley, P. L., Wendt, J. R., Vawter, G. A., Brennan, T. M. & Hammons, B. E. *Appl. Phys. Lett.* **64**, 687–689 (1994).
10. Warren, M. E. et al. *Proc. SPIE OE/LASE Conf.*, Los Angeles, 1994 (Society of Photo-instrumentation, Bellingham, Washington DC, 1994).
11. Van der Ziel, J. P. & Illegems, M. *Appl. Opt.* **14**, 2627–2629 (1975); **15**, 1256–1258 (1976).
12. Gibbs, H. M. et al. *Appl. Phys. Lett.* **41**, 221–223 (1982).
13. Gibbs, H. M. in *Optical Bistability: Controlling Light with Light* (Academic, New York, 1985).
14. Guy, D. R. P., Apsley, N., Taylor, L. L. & Bass, S. J. *SPIE* **792**, 189–193 (1987).
15. Miller, D. A. B. *IEEE J. Quantum Electron.* **QE17**, 306–313 (1981).
16. Ogura, M., Hata, T., Kawai, N. J. & Yao, T. J. *Appl. Phys. Japan* **22**, L112–L114 (1983).
17. Jewell, J. L. et al. *Electron. Lett.* **25**, 1123–1124 (1989).
18. Lee, Y. H. et al. *Electron. Lett.* **25**, 1377–1378 (1989).
19. Lee, Y. H., Tell, B., Brown-Goebeler, K., Jewell, J. L. & Hove, J. V. *Electron. Lett.* **26**, 711–712 (1990).
20. Geels, R. S., Corzine, S. W., Scott, J. W., Young, D. B. & Coldren, L. A. *Photon. Tech. Lett.* **2**, 234–236 (1990).
21. Schneider, R. P. & Lott, J. A. *Appl. Phys. Lett.* **63**, 917–919 (1993).
22. Gourley, P. L., Biefeld, R. M., Drummond, T. J. & Zipperian, T. E. *SPIE* **792**, 178–188 (1987).
23. Gourley, P. L. *Superlattices & Microstructures* **1**, 227–230 (1985).
24. Yoffe, G. W., Schlom, D. G. & Harris, J. S. Jr. *Appl. Phys. Lett.* **51**, 1876–1878 (1987).
25. Sahlen, O., Olin, U., Masseboeuf, E., Landgren, G. & Rask, M. *Appl. Phys. Lett.* **50**, 1559–1561 (1987).
26. Jewell, J. L., Scherer, A., McCall, S. L., Gossard, A. C. & English, J. H. *Appl. Phys. Lett.* **51**, 94–96 (1987).
27. Lear, K. L., Chalmers, S., Killeen, K. P. & Zolper, J. *Conf. Lasers & Electro-Optics*, paper CTu2, Baltimore, 1993 (Optical Society of America, Washington DC, 1993).
28. Peters, M. G. et al. *Proc. SPIE OE/LASE*, Vol. 2147 (Los Angeles, 1994).
29. Gourley, P. L. et al. *Appl. Phys. Lett.* **54**, 1209–1211 (1989).
30. Corzine, S. W., Geels, R. S., Scott, J. W., Yan, R.-H. & Coldren, L. A. *IEEE J. Quantum Electron.* **25**, 1513–1524 (1989).
31. Raja, M. Y. A. et al. *IEEE J. Quantum Electron.* **25**, 1513–1522 (1989).
32. Deppe, D. *Appl. Phys. Lett.* **57**, 1721–1723 (1990).
33. Schubert, E. F. et al. *Appl. Phys. Lett.* **61**, 1381–1383 (1992).
34. Osbourn, G. C. et al. in *Principles and Applications of Semiconductor Strained-Layer Superlattices, Semiconductors & Semimetals* Vol. 24 (ed. Dingle, R.) ch. 8 (Academic, London, 1987).
35. Adams, A. R. *Electron. Lett.* **22**, 249–250 (1986).
36. Yablonovitch, E. & Kane, E. O. *J. Lightwave Technol.* **LT-4**, 504–506 (1986).
37. Welch, D. F., Streifer, W., Schaus, C. F., Sun, S. & Gourley, P. L. *Appl. Phys. Lett.* **56**, 10–12 (1990).
38. Gourley, P. L., Lyo, S. K. & Dawson, L. R. *Appl. Phys. Lett.* **54**, 1397–1399 (1989).
39. Gourley, P. L. et al. *Appl. Phys. Lett.* **55**, 2698–2700 (1989).
40. Hasnain, G., Wynn, J. D., Gunapala, S. & Leibenguth, R. E. *Conf. Quantum Electron. & Laser Sci.* paper JThA2, Anaheim, 1992 (Optical Society of America, Washington DC, 1992).
41. Mukherjee, A., Mahboubzadeh, M., Schaus, C. F. & Brueck, S. R. J. *Photon. Tech. Lett.* **2**, 857–859 (1990).
42. Sinclair, M. B., Gourley, P. L., Brennan, T. M., Hammons, B. E. & Dawson, L. R. *Appl. Phys. Lett.* (submitted).
43. Gourley, P. L. *Appl. Phys. Lett.* **57**, 2410–2412 (1990).
44. Karin, H. R. et al. *Appl. Phys. Lett.* **57**, 963–965 (1990).
45. Wiesenfeld, J. M. et al. *Optical Society of America meet.*, paper FW-1 (Albuquerque, 1992).
46. Hadley, G. R. *Opt. Lett.* **15**, 1215–1217 (1990).
47. Orenstein, M., Kapon, E., Stoffel, N. G., Florez, L. & Harbison, J. P. *Conf. Lasers & Electro-Optics*, paper CPDP29-1 (Anaheim, 1990).
48. Yoo, H. J. et al. *Appl. Phys. Lett.* **56**, 1198–1200 (1990).
49. Warren, M. E. et al. *Appl. Phys. Lett.* **61**, 1484–1486 (1992).
50. Gourley, P. L., Warren, M. E., Vawter, G. A., Brennan, T. M. & Hammons, B. E. *Appl. Phys. Lett.* **60**, 2714–2716 (1992).
51. Vakhshoori, D., Wynn, J. D., Zydzik, G. J. & Leibenguth, R. E. *Appl. Phys. Lett.* **62**, 1718–1720 (1993).
52. Kittel, C. in *Introduction to Solid State Physics* 4th edn Advanced Topic A, 695–699 (Wiley, New York, 1971).
53. Yariv, A. & Yeh, P. in *Optical Waves in Crystals* ch. 4 (Wiley, New York, 1984).
54. St. J. Russell, P. J. *Appl. Phys.* **59**, 3344–3356 (1986).
55. St. J. Russell, P. J. *Physics World* 37–42 August 1992.
56. Yablonovitch, E. *Phys. Rev. Lett.* **58**, 2059–2062 (1987).
57. Ho, K. M., Chan, C. T., Soukoulis, C. M. *Phys. Rev. Lett.* **65**, 3152–3155 (1990).
58. John, S. *Physics Today* **44**, 32–32 (1991).
59. Haus, J. W. *J. mod. Optics* **41**, 195–207 (1994).
60. Sözüer, H. S. & Dowling, J. P. *J. mod. Optics* **41**, 231–239 (1994).
61. Robertson, W. et al. *Phys. Rev. Lett.* **68**, 2023–2026 (1992).
62. Özbay, E. et al. *Phys. Rev. Lett.* **64**, 2059–2061 (1994).
63. Ho, K. M., Chan, C. T., Soukoulis, C. M., Biswas, R. & Sigalas, M. *Solid State Commun.* **89**, 413–416 (1994).
64. Meade, R. D., Brommer, K. D., Rappe, A. M. & Joannopoulos, J. D. *Appl. Phys. Lett.* **61**, 495–497 (1992).
65. Villeneuve, P. R. & Piche, M. *Proc. Quantum Electronics Laser Science Conf.* paper QWG6, Anaheim, 1992 (Optical Society of America, Washington DC, 1992).
66. Yamamoto, Y. & Björk, G. J. *Appl. Phys. Japan* **30**, L2039–L2044 (1991).
67. Faist, J., Capasso, F., Sivco, D. L., Sirtori, C., Hutchinson, A. L. & Cho, A. Y. *Science* **264**, 553–556 (1994).
68. Tsu, R. *Nature* **369**, 442–443 (1994).
69. Morgan, R. A. *Proc. SPIE OE/LASE Conf.*, Los Angeles, 1994 (Society of Photo-instrumentation, Bellingham, Washington DC, 1994).

ACKNOWLEDGEMENTS. I thank colleagues at Sandia National Laboratories who have collaborated on microstructured lasers and materials, including T. Brennan, W. Chow, R. Dawson, T. Drummond, G. Hammons, R. Hadley, K. Lear, A. McDonald, R. Schneider, M. Sinclair, A. Vawter, M. Warren, J. Wendt & T. Zipperian. I also thank R. Slusher of AT&T Bell Laboratories and R. Morgan of Honeywell for supplying photographs of Fig. 5 and Fig. 6, respectively. The work was supported in part by the Department of Energy and the Office of Basic Energy Sciences.



# The crystal structure of the bacterial chaperonin GroEL at 2.8 Å

Kerstin Braig<sup>\*†</sup>, Zbyszek Otwinowski<sup>†‡§</sup>, Rashmi Hegde<sup>†‡§</sup>, David C. Boisvert<sup>\*</sup>, Andrzej Joachimiak<sup>‡§</sup>, Arthur L. Horwich<sup>\*†</sup> & Paul B. Sigler<sup>†‡||</sup>

<sup>\*</sup> Department of Genetics and <sup>‡</sup> Department of Molecular Biophysics and Biochemistry and <sup>†</sup> Howard Hughes Medical Institute, Yale University School of Medicine, Boyer Center, 295 Congress Avenue, New Haven, Connecticut 06510, USA

**The crystal structure of *Escherichia coli* GroEL shows a porous cylinder of 14 subunits made of two nearly 7-fold rotationally symmetrical rings stacked back-to-back with dyad symmetry. The subunits consist of three domains: a large equatorial domain that forms the foundation of the assembly at its waist and holds the rings together; a large loosely structured apical domain that forms the ends of the cylinder; and a small slender intermediate domain that connects the two, creating side windows. The three-dimensional structure places most of the mutationally defined functional sites on the channel walls and its outward invaginations, and at the ends of the cylinder.**

CHAPERONINS are large multisubunit assemblies essential in mediating ATP-dependent polypeptide chain folding in a variety of cellular compartments (for review see refs 1–4). Two families of chaperonins have been identified, the Hsp60 class, with members in the bacterial cytoplasm (GroEL) and in endosymbiotically derived mitochondria (Hsp60) and chloroplasts (Rubisco binding protein); and the TF55/TCPI family in thermophilic archaea and the evolutionarily connected eukaryotic cytosol. Members of both families consist of two stacked rings<sup>5,9</sup>, each ring containing radially arranged subunits of relative molecular mass ~60,000 ( $M_r$ , ~60K), with seven identical subunits per ring in the GroEL/Hsp60 family<sup>5–7</sup>. Even though some members of each family can be heat-induced, genetic studies indicate that chaperonins are essential in protein folding at all temperatures<sup>10–13</sup>. When chaperonin function is deficient, newly translocated or newly synthesized proteins fail to reach native form and are liable to aggregate<sup>11,12</sup>.

*In vitro* studies with polypeptides diluted from denaturant show that non-native folding intermediates bind to GroEL with a stoichiometry of one or two polypeptides per GroEL 14-mer<sup>14–17</sup>. Binding must occur through a common feature that is shared by many polypeptide folding intermediates, but is not present in native forms<sup>18–21</sup>. The exact site on GroEL of polypeptide binding is not known, although electron microscopic studies indicate that part of the bound polypeptide occupies the central cavity of the rings<sup>22–25</sup>. Efficient release and folding requires the hydrolytic turnover of ATP<sup>17,26–28</sup> and, for many polypeptides, the cooperating action of GroES, a single ring of seven identical 10K subunits that binds at one or both termini of the GroEL cylinder<sup>22,24,29–33</sup>. Chaperonins appear to act by stabilizing non-native intermediates against off-pathway misfolding and aggregation by a mechanism that is still unknown<sup>1–4</sup>. To provide the required stereochemical framework for further genetic, biochemical and biophysical studies aimed at understanding chaperonin-assisted protein folding, we have determined the three-dimensional structure of unliganded GroEL by X-ray crystallography to 2.8 Å.

## Crystallization and structure determination

GroEL was overexpressed in *E. coli* from an inducible *trc* promoter and purified as described in the legend to Table 1. The protein that produced the best crystals carries two mutational changes, R13G and A126V. This variant appears to function normally, as expression from the cloned operon rescues a GroEL-deficient strain, and the purified complex exhibits normal binding and GroES-mediated folding *in vitro* (M. Goldberg, D.C.B., A.L.H., data not shown). Orthorhombic crystals, space group C22<sub>1</sub> ( $a = 178$ ,  $b = 203$ ,  $c = 278$  Å), were grown by vapour diffusion as described in the legend to Table 1. The crystals diffract to at least 2.7 Å (Table 1) and contain 7 subunits per asymmetric unit. Self-rotation (MERLOT<sup>34</sup>) using data to 4 Å, indicated an unambiguous 7-fold rotation axis nearly parallel to the  $c$  axis and perpendicular to a dyad, implying that the rings were 7-fold rotationally symmetrical and stacked 'back-to-back' with a dyad between them (Fig. 1). This architecture is consistent with previous electron microscopy and crystallographic studies<sup>5,24,51</sup>. Initial phases were obtained from a single isomorphous ethyl mercuric derivative. The first heavy-atom site was located by a 6-parameter search that maximized the correlation at 6 Å resolution between heavy-atom-induced intensity differences  $\Delta F_H^2$  and the calculated heavy-atom structure factors  $|F_H^C|^2$ . The six parameters were: the position of the 7-fold axis of rotational symmetry in the  $x$ - $y$  plane (two highly restricted parameters because the 7-fold axis was assumed to intersect a lattice dyad), the position of the heavy atom in the reference subunit (three parameters, expanded to seven sites by 7-fold rotational symmetry), and the angle by which the 7-fold rotation axis deviated from the  $c$  axis (one parameter assuming the 7-fold axis is perpendicular to a lattice dyad). Once the provisional 'first site' was located and fixed, a second site and its occupancy relative to the first site was established using the same procedure. Similarly a third site and its occupancy was established. A similar search for a fourth site and difference Fourier analysis after heavy-atom refinement revealed no further sites. Twenty-one sites (3 sites  $\times$  7-fold symmetry) were refined with MLphare<sup>35</sup> with no local symmetry constraints (Table 1, derivative refinement). The anomalous scattering differences were not used. The results of this heavy-atom refinement are shown in Table 1. Applying the 7-fold rotation deduced in the heavy-atom search procedure (the 7-fold axis was 1.5° from  $c$ ) caused the heavy-atom sites, which had been refined without rotational constraints, to superimpose with an r.m.s. deviation of 0.5 Å. The

§ Present addresses: Department of Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, Texas 75235-9038, USA (Z.O.); Skirball Institute of Biomolecular Medicine, New York University Medical Center, New York, New York 10016, USA (R.H.); Argonne National Laboratory, 9700 Cass Avenue, Argonne, Illinois 60439-4833, USA (A.J.).

|| To whom correspondence should be addressed.

electron-density map phased with a single isomorphous derivative was not interpretable, but when averaged around the 7-fold rotation axis a clear molecular boundary was evident for the assembly, as were the rough outlines of the subunits. Starting at 6 Å and proceeding to 2.7 Å, 400 cycles of phase improvement/phase extension were done by averaging about the same 7-fold axis. A new envelope was assigned automatically every 20 cycles using the characteristic fluctuation in electron density as a criterion of molecular density. Details of this process will be published elsewhere. Maximum entropy weighting was applied to minimize the difference between calculated density-modified and observed amplitudes (Z.O., manuscript in preparation). The result was a combined single isomorphous replacement (SIR)/averaged map whose side chain and backbone density was readily interpretable in the equatorial and intermediate domains but poorly defined in parts of the apical domain (see below).

The model derived from the SIR/symmetry-averaged density ( $M^{\text{SIR/Av}}$ ) was not used to start the refinement but rather to initiate a second independent phasing procedure using more accurate higher resolution data collected at the EMBL outstation at DESY (Table 1). Spheres of 5 Å radius were centred on the C $\alpha$  positions of  $M^{\text{SIR/Av}}$  to produce a starting subunit envelope. Random phases were applied to the amplitudes from 30 Å to 8 Å resolution and the electron-density map subjected initially to non-crystallographic symmetry (NCS) averaging and solvent

flattening using the 7-fold rotation axis and molecular envelope derived from  $M^{\text{SIR/Av}}$ . Phases were gradually extended to higher resolution amplitudes (ultimately to 2.7 Å) with periodic adjustments of the NCS operators using the program RAVE (G. L. Kleywegt and T. A. Jones, personal communication; and ref. 36) which uses the CCP4 averaging algorithm. In RAVE, the original 7-fold rotational symmetry operator was replaced after the first cycle by a series of general superimposition matrices. The resulting electron density (such as in Fig. 2b) had the same backbone connectivity as  $M^{\text{SIR/Av}}$  but significantly improved detail in the previously ambiguous segments of the apical domain.

## Modelling and refinement

The electron density map (Fig. 2b) was easily fitted with a skeletal model by the program O<sup>37</sup> even though segments of electron density in the apical region facing the central channel and toward the top surface were either locally fragmented and/or showed poor side-chain density. As will be discussed below, some of these segments have been implicated by mutagenesis<sup>38</sup> to be involved in peptide binding and could be inherently flexible and poorly ordered for functional reasons. The amino-terminal 5 residues and the carboxy-terminal 24 residues were not visualized. The model was refined to 2.8 Å with X-PLOR<sup>39</sup> under NCS restraints starting with 4 Å rigid body adjustments of three

TABLE 1 Crystallographic data

|                                                               |                 |                 |             |             |             |             |             |             |             |             |                |
|---------------------------------------------------------------|-----------------|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|----------------|
| Parent data                                                   |                 |                 |             |             |             |             |             |             |             |             |                |
| <b>Resolution shells* (Å)</b>                                 | <b>&gt;5.80</b> | <b>4.61</b>     | <b>4.03</b> | <b>3.66</b> | <b>3.40</b> | <b>3.20</b> | <b>3.04</b> | <b>2.91</b> | <b>2.80</b> | <b>2.70</b> | <b>Overall</b> |
| Coverage (%)                                                  | 99.9            | 99.6            | 98.8        | 97.2        | 92.9        | 87.6        | 84.8        | 83.8        | 82.5        | 53.6        | 88.2           |
| $\langle I/\sigma \rangle^\dagger$                            | 27.2            | 21.8            | 20.0        | 14.3        | 10.1        | 7.0         | 4.8         | 3.5         | 2.7         | 1.6         | 12.4           |
| Derivative refinement                                         |                 |                 |             |             |             |             |             |             |             |             |                |
| <b>Resolution shells (Å)</b>                                  |                 | <b>&gt;16.1</b> | <b>11.1</b> | <b>8.4</b>  | <b>6.8</b>  | <b>5.7</b>  | <b>4.9</b>  | <b>4.3</b>  | <b>3.8</b>  |             | <b>Overall</b> |
| Phasing power <sup>‡</sup> of Et-Hg <sub>1</sub> <sup>§</sup> |                 | 2.17            | 2.68        | 2.03        | 1.83        | 1.45        | 1.02        |             |             |             | 1.68           |
| Phasing power of Et-Hg <sub>2</sub>                           |                 | 1.74            | 2.59        | 2.04        | 2.01        | 1.74        | 1.28        | 0.93        | 0.69        |             | 1.27           |
| Model refinement                                              |                 |                 |             |             |             |             |             |             |             |             |                |
| <b>Resolution shells (Å)</b>                                  | <b>5.68</b>     | <b>4.66</b>     | <b>4.11</b> | <b>3.76</b> | <b>3.50</b> | <b>3.30</b> | <b>3.14</b> | <b>3.01</b> | <b>2.90</b> | <b>2.80</b> | <b>Overall</b> |
| Working R-factor¶                                             | 0.366           | 0.313           | 0.279       | 0.294       | 0.314       | 0.329       | 0.348       | 0.367       | 0.389       | 0.393       | 0.326          |
| Free R-factor¶                                                | 0.424           | 0.369           | 0.320       | 0.343       | 0.358       | 0.362       | 0.369       | 0.411       | 0.400       | 0.403       | 0.368          |

B-factors were held uniform over each domain: 18.8 Å<sup>2</sup> for equatorial and intermediate domains; 47.5 Å<sup>2</sup> for the apical domain; the r.m.s. deviation of bond distances and interbond angles is 0.006 Å and 1.12°, respectively. GroEL was overexpressed in *E. coli* DH5 from a plasmid joining the GroE operon with a *trc* promoter. The cells were induced with 1 mM isopropyl- $\beta$ -D-thiogalactoside (IPTG) at 37 °C when the A<sub>650</sub> reached 0.4. After 2 h, cells were collected in a buffer containing 50 mM Tris-HCl (pH 7.4), 1 mM EDTA, and 1 mM DTT, and broken at 4 °C in a microfluidizer (Microfluidics Corp.). The lysate was clarified by centrifugation at 100,000g for 1 h, loaded onto a Q-Sepharose Fastflow column (Pharmacia), and the column was eluted with a linear gradient of 0–400 mM NaCl containing the lysis buffer components. The GroEL-containing fraction (~400 mM NaCl) was concentrated to ~50 mg ml<sup>-1</sup> in a Centricon 100 and washed against 20 mM Tris-acetate pH 8.0. Crystals were obtained from hanging drops (4  $\mu$ l) initially containing 25 mg ml<sup>-1</sup> GroEL, 0.2% (w/v) PEG 8000, 0.86 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2 mM CaCl<sub>2</sub>, 50 mM Tris-acetate (pH 8.0), equilibrated at 28 °C against reservoirs of 0.4% (w/v) PEG 8000, 1.72 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 4 mM CaCl<sub>2</sub>, 100 mM Tris-acetate (pH 8.0). Crystals appeared overnight and grew typically to 0.5 mm  $\times$  0.5 mm  $\times$  0.1 mm in 2 weeks. Before diffraction studies or heavy-atom soaks, the crystals were stabilized in 1.8 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.4% PEG 8000, 4 mM CaCl<sub>2</sub>, 20 mM Tris-acetate (pH 8.0) and 20% (w/v) glycerol. The crystals belong to the space group C222<sub>1</sub>, with  $a = 178$  Å,  $b = 204$  Å,  $c = 278$  Å, and diffract isotropically to 2.7 Å. The volume of the asymmetric unit,  $1.26 \times 10^6$  Å<sup>3</sup>, can accommodate 7 subunits of 58K ( $V_m = 3.1$  Å<sup>3</sup> per unit  $M_r$ ) but not 14 ( $1.55$  Å<sup>3</sup> per unit  $M_r$ ). Derivatives were prepared by soaking crystals for 24 h in stabilizer containing 0.5 mM ethylmercuric chloride or 0.5 mM methylmercuric chloride. Diffraction data were collected in a N<sub>2</sub> stream at between -150° and -160 °C. Crystals were frozen by suspending them in a film of stabilizing supernatant in nylon loops and plunging into a melting mixture of solid/liquid propane. Crystals encased in solid propane were placed in a N<sub>2</sub> stream (-150 °C) and the diffracting position exposed once the propane had dissipated. Parent data were collected on an R-axis II imaging plate detector (mirror focused CuK $\alpha$  X-rays) and on a MaResearch imaging plate detector at the EMBL outstation at DESY ( $\lambda = 0.92$  Å). The ethylmercuric and methylmercuric data were collected on a Siemens-Xentronic detector (mirror focused CuK $\alpha$ ) and on imaging plates/Fuji scanner at the F1 beam line at CHESS ( $\lambda = 0.91$  Å). All data were processed and merged with DENZO<sup>50</sup> and SCALEPACK (Z. Otwinowski).

\* Shells represent the regions of the diffraction pattern that have resolutions greater than the value of column heading and smaller than the heading of the column to the left.

† Measure of the statistical reliability of the intensity in the shell; the average value of the intensity relative to the counting statistics error.

‡ The magnitude of the heavy-atom's effect on the intensity normalized to the estimated error in the phase analysis. Phasing power is  $\sum_h |F_h^o| / \sum_h |F_h^c| \exp(i\Phi_h^c) + F_h^o - |F_h^c|$ , where  $F_h^o$  = heavy-atom structure factor and  $|F_h^c|$  and  $\Phi_h^c$  are observed amplitudes for the protein and heavy-atom derivative, respectively,  $\Phi_h^c$  is the calculated phase, and the sum is over all observations.

§ Three sites per subunit; two separate data sets.

|| 58 residues, including those indicated in Fig. 2c as questionable, were omitted.

¶ Free R-factor calculated with a random 10% of the intensities evenly distributed throughout the resolution range, working R-factor calculated with the remaining intensities.

The refined coordinates for alpha carbons will be deposited in the Protein Data Bank (Brookhaven).



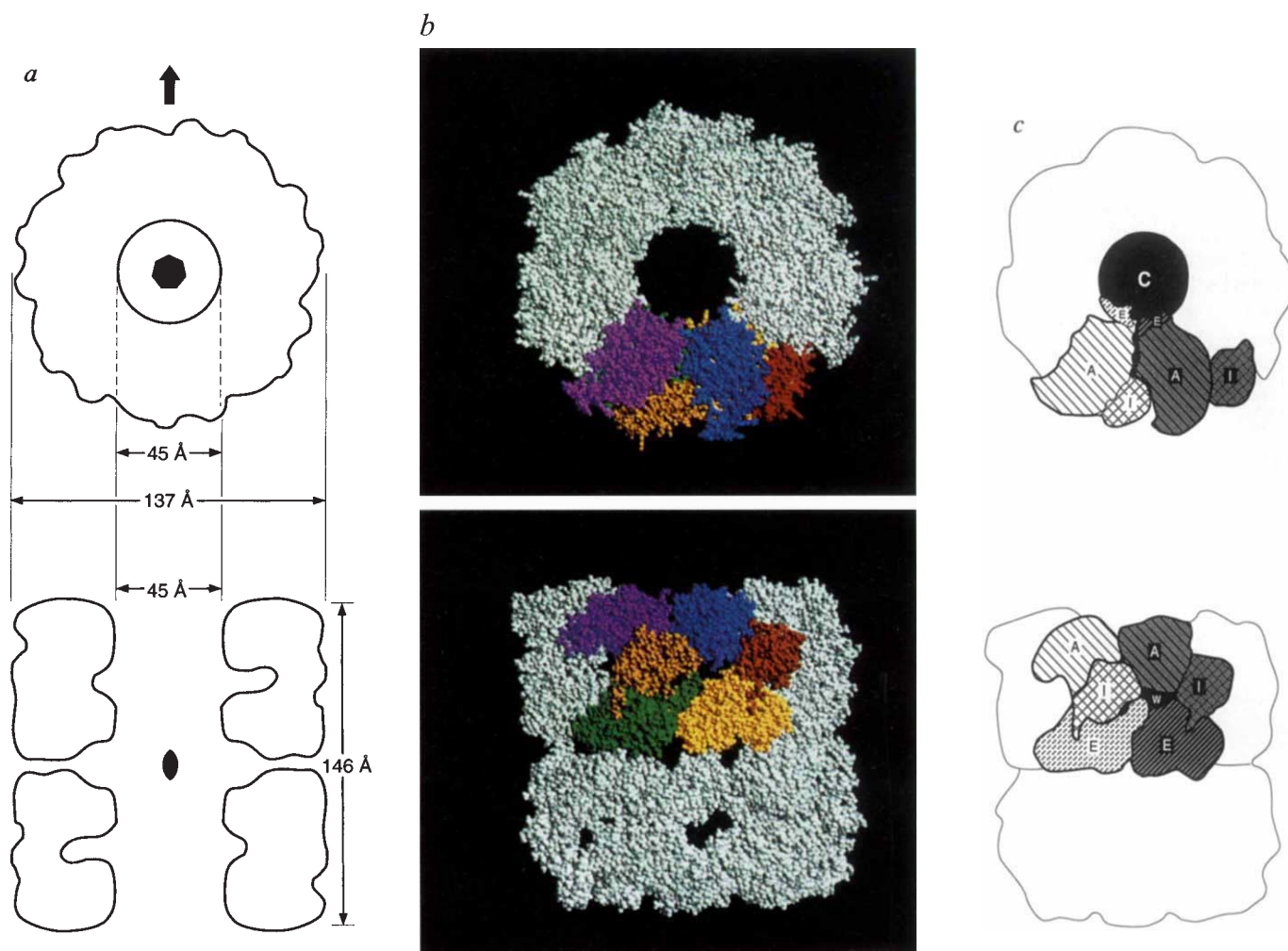


FIG. 1 Architecture of GroEL. *a*, Principal dimensions on a schematic diagram of the double-ring cylinder, showing a view into the central channel (top panel), and a vertical cut through the cylinder made by a plane defined by a diameter and the cylindrical axis (bottom panel). The dimensions are rough because the surfaces are irregular. The walls of the inner channel contain substantial bulges at the level of the intermediate domain which are not shown. *b*, Van der Waals space-filling model of the entire 14-mer shown in grey except for two adjacent subunits in the 'top' ring seen facing into the channel (top panel) and

viewed from the outside (bottom panel). Domains within the left subunit are colour-coded as follows: equatorial, green; intermediate, gold; and apical, purple; the right subunit's equatorial, intermediate and apical domains are colour-coded yellow, red and blue, respectively (see Figs 2, 3 and text for discussion of domains). *c*, Drawing depicting diagrammatically the structure in *b*. E, I and A, equatorial, intermediate and apical domains, respectively; C, central channel; W, external opening of a side window.

distinct domains (see below), followed by positional refinement and simulated annealing to 2.8 Å. Most of the model was easily adjusted with the appropriate SIGMA-A-phased<sup>40</sup> difference maps. However, these maps did not improve the discontinuous/ambiguous segments noted above, namely residues 1–5, 207–210, 225–246, 255–275, 303–314, 525–547 (see Fig. 2*c* for details). Indeed, the best *R*-factors were obtained when segments containing these residues were deleted and 'omit' refinements/maps failed to yield interpretable density in these regions. There was no significant improvement in the refinement when the NCS constraints were released, either in total or just in the troublesome apical domain. Although we cannot rule out misconnected density on the basis of this refinement alone, our interpretation is supported by contiguous density in the ATP-bound form of GroEL (D.C.B., J. Wang, Z.O., A.L.H. and P.B.S., unpublished results).

Figure 2*c* shows an assessment of the reliability of the refined structure that is shown in the illustrations throughout this paper and the companion paper on mutagenesis<sup>38</sup>. The Bowie–Eisenberg three-dimensional profile<sup>41</sup> (Fig. 2*c*) shows an excellent

overall score, indeed comparable to that for well-refined 2.0 Å structures; however, the three-dimensional–one-dimensional plot shows a distinct valley indicating that the structure is likely to be incorrect in a region described above as being questionable (297–319). Figure 2*c* shows that this region also corresponds to one of the regions with poorly interpretable electron density. Finally, the bar graph displays significant deviations from acceptable backbone torsion angles. Here again at 297–319 there are serious  $\phi, \psi$  violations; indeed, 11 of the 17 violations overall occur in regions where the interpretation is not supported by electron density. Thus the segments on the surface of the apical domain (totalling 50 residues), denoted by shaded and solid bars in Fig. 2*c*, are not to be considered as experimentally well defined conformations. It is not clear to what extent these uncertainties reflect true conformation variability. A substantial improvement in the amount and accuracy of high-resolution data will be required before this can be resolved. Note that, except for the extreme amino- and carboxy-termini, the troublesome segments occur only in the apical domain, where mutagenesis has identified the binding sites for unfolded polypeptide. Clearly a degree

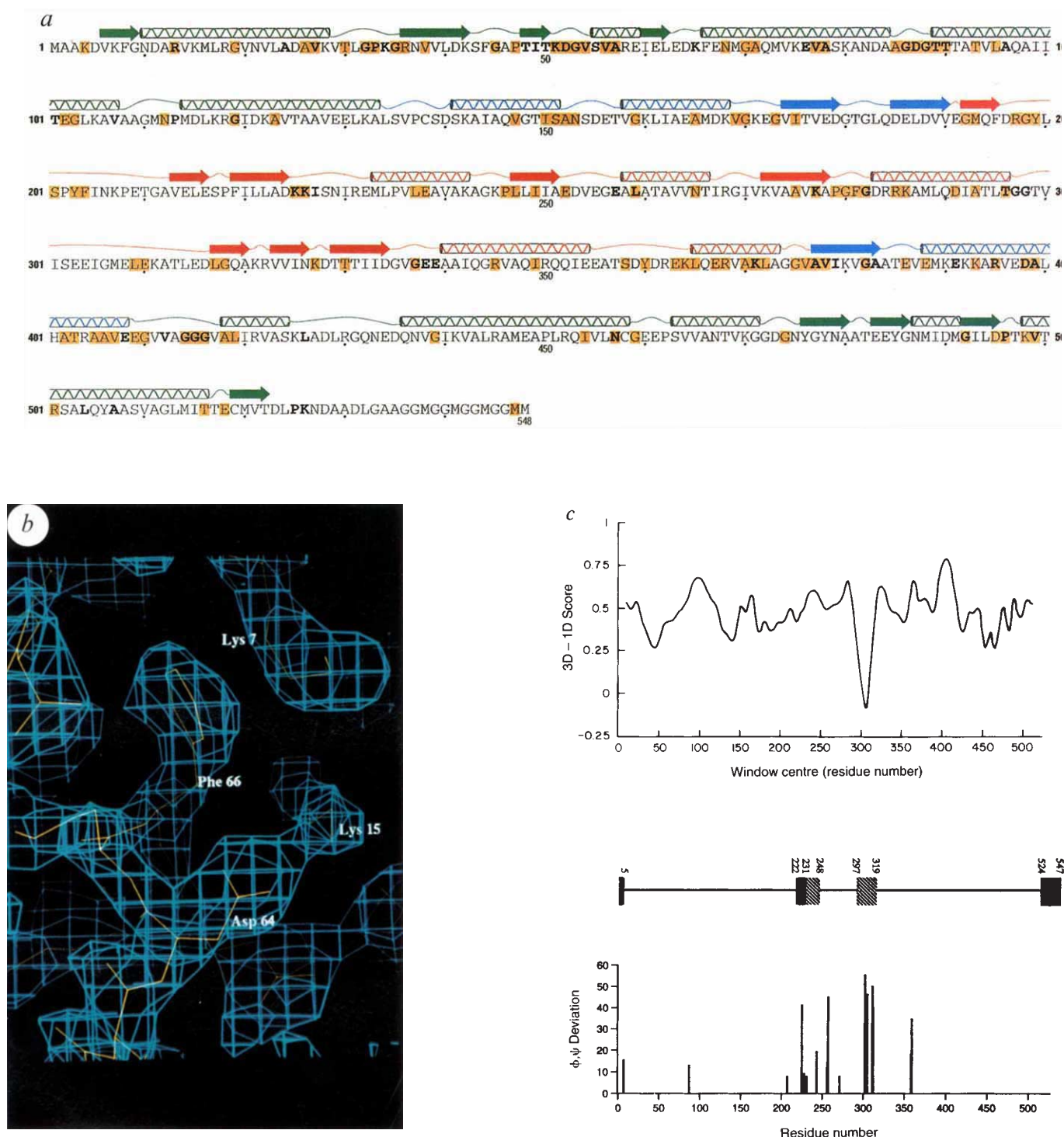


FIG. 2 Structural components of a subunit. **a**, Amino-acid sequence. Secondary structural elements (defined by visual inspection), indicated by arrows ( $\beta$ -strands) or sine wave in rectangles ( $\alpha$ -helices), and extended strands. Colour-coding corresponds to the alpha carbon trace in Fig. 3c (not to Fig. 1) and denotes in which domain the sequence segment occurs; here, equatorial is green, intermediate blue, apical red. Residues that are greater than 98% conserved in GroEL/Hsp60 chaperonins are in orange-coloured boxes; those conserved also in TF55/TCP1 are in bold. **b**, Electron density,  $(2|F^o| - |F^c|)$  exp ( $i\Phi^c$ ), contoured at  $2.2\sigma$  with the refined model superimposed.  $|F^o|$ ,  $|F^c|$  and  $\Phi^c$  are the observed and calculated amplitudes and phases calculated from the refined model. **c**, Reliability of the refined model. Top panel is plot of Bowie-Eisenberg score (3D-1D)<sup>41</sup>, which indicates the conformity of the environment for a 7-residue segment centred on the designated residue with that predicted by a consensus of 200 well refined

high-resolution structures ( $d \leq 2.0$  Å) from the protein data bank. Scores below 0.2 indicate poor agreement with expectation, below 0.1 suggest the structure might be wrong. Notwithstanding the valley at 300, the overall score is 248, which is expected for a well refined structure at 2.0 Å. The middle panel denotes regions that are poorly modelled. The crosshatched segments have marginally contiguous density for the backbone and uninterpretable side-chain density in the experimental map. The solid black bars indicate badly discontinuous or absent density. These regions were not improved with  $(2F^o - F^c)$  maps or omit maps. Except for the N-terminal five residues and C-terminal 24 residues the questionable regions were modelled with the program O<sup>37</sup>. Bottom panel shows a bar plot of the deviations from the normal backbone angles  $\phi, \psi$ . The y axis indicates the shortest distance on the  $\phi, \psi$  plane (greater than  $6^\circ$ ) to the boundary of acceptable regions.



of flexibility is likely in structural components designed to bind a wide range of polypeptide sequences.

### General architecture

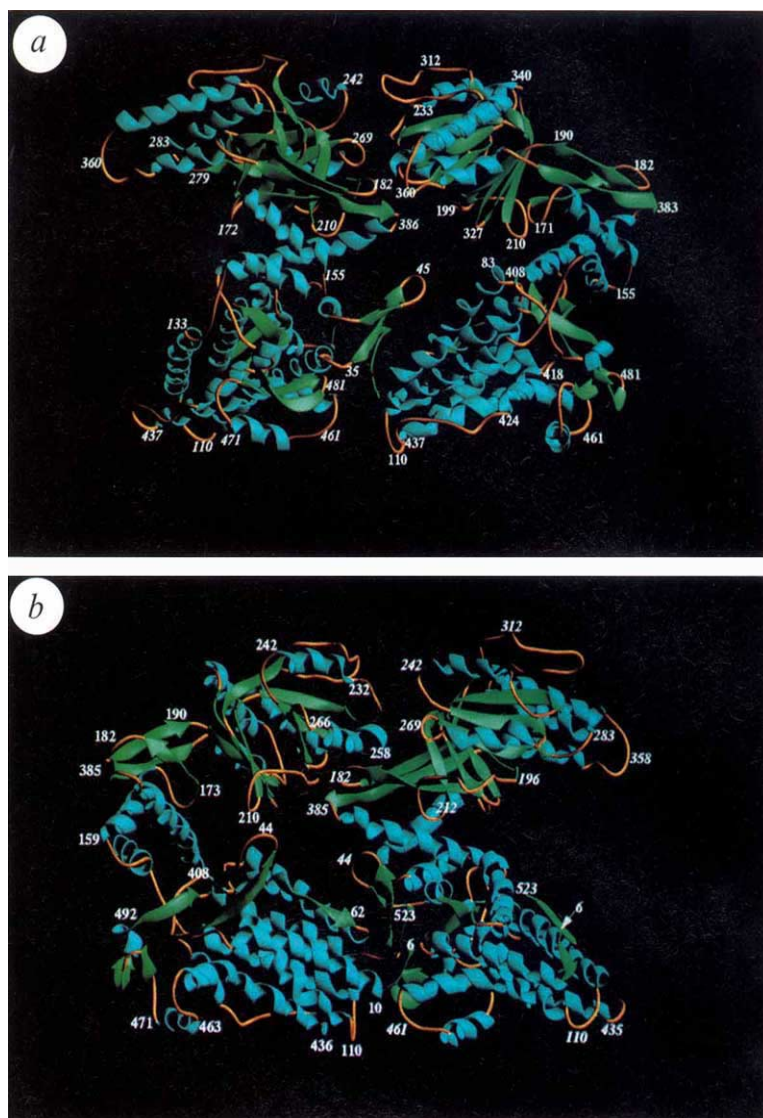
GroEL is a porous thick-walled cylinder that is slightly taller than it is wide and contains a substantial central cavity or channel. It is composed of two rings of seven large subunits (547 amino acids) arranged with nearly exactly 7-fold rotational symmetry. The rings are arranged back-to-back forming an extensive interface with one another across a remarkably flat equatorial plane. The plane contains a crystallographic dyad which is propagated by the 7-fold rotation axis into seven molecular dyads in the equatorial plane. The overall dimensions are shown in Fig. 1a.

### Domain structure of the subunits

The subunits are folded into three distinct domains that are arranged as shown in Fig. 1b, c. The residues that make up the domains are indicated in Fig. 2a and their structures are shown in Fig. 3c, d. The equatorial domain, the largest domain (residues 6–133 plus 409–523, totalling 243 residues), contains the segments adjacent to the crystallographically disordered N-terminal and C-terminal residues, which appear to project into the central channel. This domain is highly  $\alpha$ -helical and well ordered. It serves as the foundation of the chaperonin structure providing most of the side-to-side contacts between subunits in the ring, and all of the contacts between the rings across the equatorial plane (Table 2 and Fig. 4b). In addition to forming a solid wall around the equatorial waist of the cylinder, it provides the ATP binding site established by mutational analysis<sup>38</sup> and visualized in detail in the crystal structure of the ATP-bound form (D.C.B., J. Wang, Z.O., A.L.H. and P.B.S., unpublished results).

The intermediate domain is the smallest (residues 134–190 plus 377–408, totalling 89 residues) and provides a covalent connection from the equatorial to the apical domain. The secondary structural elements of the intermediate domain form a diagonal projection which contacts the apical domain of the adjacent subunit on the right (unless otherwise stated, all descriptions refer to the 'top' ring with the observer on the outside as in Fig. 1b). This projection also forms the left wall and roof of an ellipsoidal window or portal (described in more detail below) which penetrates through the wall of the cylinder into the central channel. The base of the intermediate domain forms an acute angle with the equatorial domain and flanks the ATP binding site. The intermediate domain forms part of the outer circumference of the cylinder. As it extends only  $\sim 26$  Å in the radial direction, the wall of the cylinder is thin at this point, causing a bulging out of the central cavity to its widest point ( $\sim 90$  Å). The covalent connections of this domain to the equatorial domain and apical domain are through short antiparallel segments which could easily serve as hinges in allosteric adjustments.

The apical domain (residues 191–376, 186 residues) forms the opening of the central channel. Its thermal parameters, on average, are much higher than the rest of the structure and vary considerably within the domain. The portions facing away from the channel and those in contact with the intermediate domain to the left are reasonably well ordered. In contrast, some of the segments facing the channel and the top surface have very high temperature factors and segments identified above and shown in Fig. 2c cannot be visualized on omit refinement/maps phased with the remaining 90% of the structure. As described in ref. 38, these segments bear mutational changes that disrupt polypeptide and GroES binding, suggesting that these segments may be inherently flexible. It is possible that the apical domain exhibits



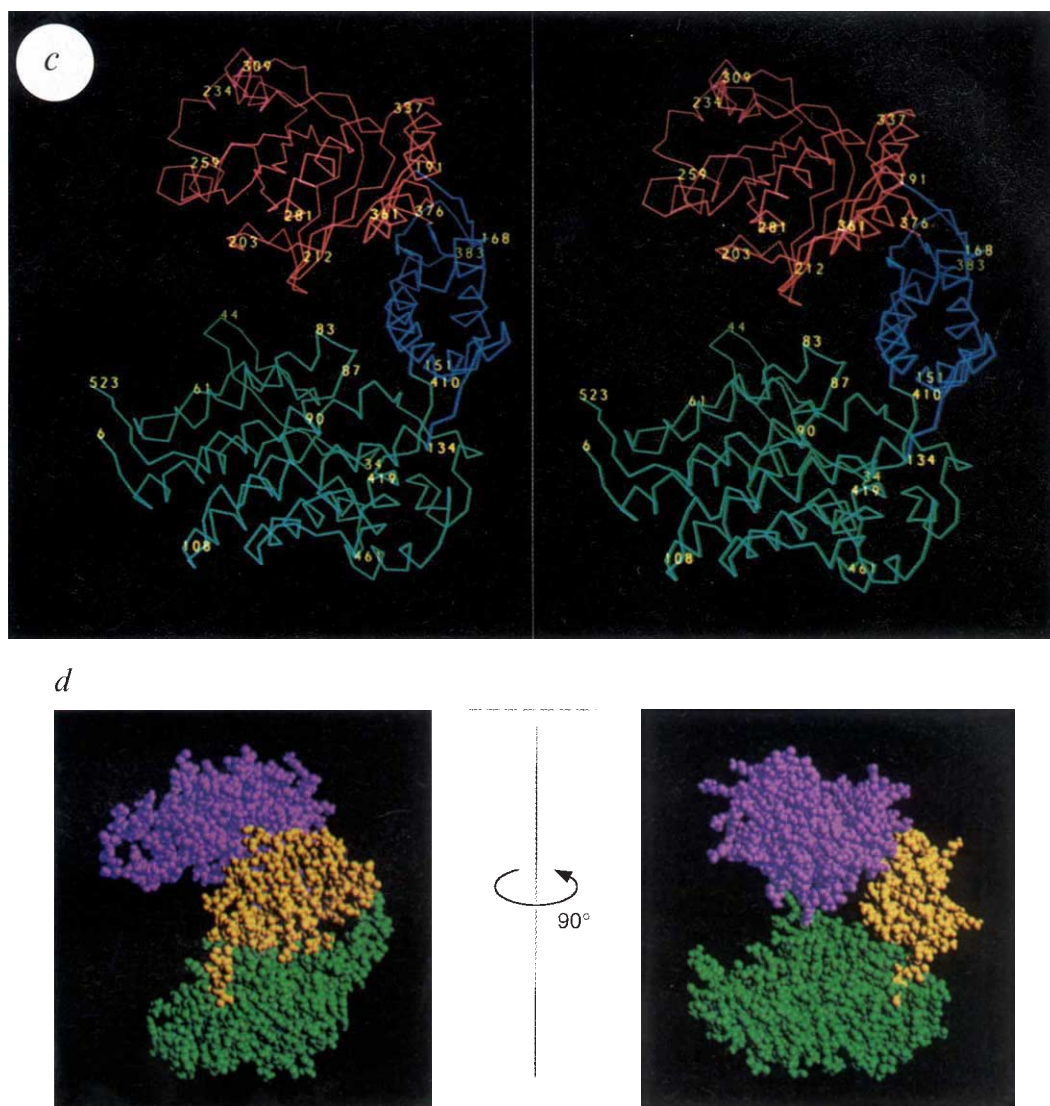
two forms of flexibility: an *en-bloc* movement generated by hinge-like movements at its union with the intermediate domain, and a local or segmental flexibility enabling these regions to complement the wide range of unfolded polypeptide 'substrate' sequences which bind in this region.

### Contact between subunits

Each subunit contributes 6.6% of its solvent-accessible surface<sup>42</sup> (or 1,400 Å<sup>2</sup>) to intersubunit contacts, an area that is typical of specific protein-protein interactions<sup>43</sup> (for example 1,600 ± 350 Å<sup>2</sup> for antibody-antigen interaction), but it is at the low end of the range expected for intersubunit contacts in oligomeric proteins<sup>44,45</sup>. This leaves a substantial 93.4% of surface, or 19,800 Å<sup>2</sup> per subunit, for interaction with solvent components, nucleotides, unfolded/partially folded peptides and GroES. For each subunit, side-by-side contacts in the ring contribute 982 Å<sup>2</sup> and the remaining 418 Å<sup>2</sup> of buried surface involves contacts across the equatorial plane. The solvent-accessible surface of an isolated subunit is 57.9% nonpolar which is only 1% greater than the average for globular proteins and well within the r.m.s. deviation of 5%. The surface made inaccessible by intersubunit contacts is 59.0% nonpolar which is noticeably less than the average value of 67% for the subunit interfaces of oligomeric proteins. The equatorial domain does not contact the intermediate or apical domain of its neighbouring subunit.

There are two notable features to the equatorial domain contacts as shown in Fig. 4a. An antiparallel  $\beta$ -loop (residues 36–

FIG. 3 Conformations within a subunit. *a*, Ribbon drawing of two adjacent subunits in the top ring (see Fig. 1) viewed from the outside. Beta-strands are wide green arrows, alpha-helices are blue, and connecting strands are yellow tubes. *b*, The same two subunits viewed from the inside of the channel. *c*, Stereopair of an alpha carbon trace colour-coded as in Fig. 2*a*. *d*, Two orthogonal views of the space-filling model of the left subunit of Fig. 1*b*. Left panel as in Fig. 1*b*; the right panel shows the view in the left panel rotated 90° around a vertical axis so that the 'channel-facing' surface is on the left and the 'outside' surface is on the right. Backbone elements of the equatorial and apical domain that appear to touch are greater than 5 Å apart but could be easily bridged by water molecules or brought into contact by small conformational shifts in either or both domains.



50) projects from the body of the domain towards the inner surface of its right-handed neighbour where it forms a parallel  $\beta$ -structure near its neighbour's C-terminal segment (519–522). This near C-terminal segment is supported by an anti-parallel  $\beta$ -strand interaction with the near N-terminal segment (6–8) of the same subunit. Deletion mutagenesis supports the role of this intersubunit interaction in assembly: although a C-terminal deletion starting at residue 522 is tolerated, deletions including residue 521 are lethal, associated with impaired GroEL assembly<sup>52</sup>. The equatorial plane between the rings (Fig. 4*b*) is an extensive smooth van der Waals contact surface. Although both polar and nonpolar components contribute to this surface, there are no loops or even large side chains penetrating through the plane. A list of residues contributing to intersubunit contacts is given in Table 2.

As noted earlier, the upper right corner of the intermediate domain, as it arches over the portal on the outside of the GroEL cylinder, makes contact with the bottom of the apical domain in the neighbouring subunit to the right. Thus, the intermediate domain provides a scaffold for potential allosteric communication between the equatorial and apical domains both within the subunit and through non-covalent contact between subunits (the latter connecting the equatorial/intermediate domain on the left with the apical domain on the right). Unlike the seemingly flexible segments on the inner surface of the apical domain, the segments contacting the neighbouring intermediate domain involve well defined structures, suggesting an important role in

assembly and possibly in the modulation of function. A stereochemical description of the role of the various domains in the allosteric behaviour of GroEL awaits the completion of structure determinations of the appropriate complexes as well as correlated mutagenesis and biochemical experiments.

### The central channel

A long-recognized striking feature of GroEL is a large central cavity that, in the crystal structure, appears to traverse the entire length of the cylinder with no obstruction. This contrasts with electron microscopy reconstructions which show density completely obstructing the channel at both the apical and equatorial levels of each ring<sup>24</sup>. Given conclusions about the function of the apical domain in peptide binding (see ref. 38), the electron microscopy densities near the channel opening could be derived from bound polypeptide. From the position and orientation of contiguous sequence segments, the disordered N and C termini of the GroEL subunits would project into the channel from the equatorial domain; thus the equatorial densities seen in the electron microscopy may result from fixation of the C-terminal 25 residues, which are not seen in the crystal structure presumably because they are either flexible or fixed in a variety of conformations.

Except for one loop projecting into the channel, the diameter of the channel is quite uniform, averaging about 47 Å. There is a considerable outward bulge in each ring at the level of the intermediate domain. The channel's total length (continuous



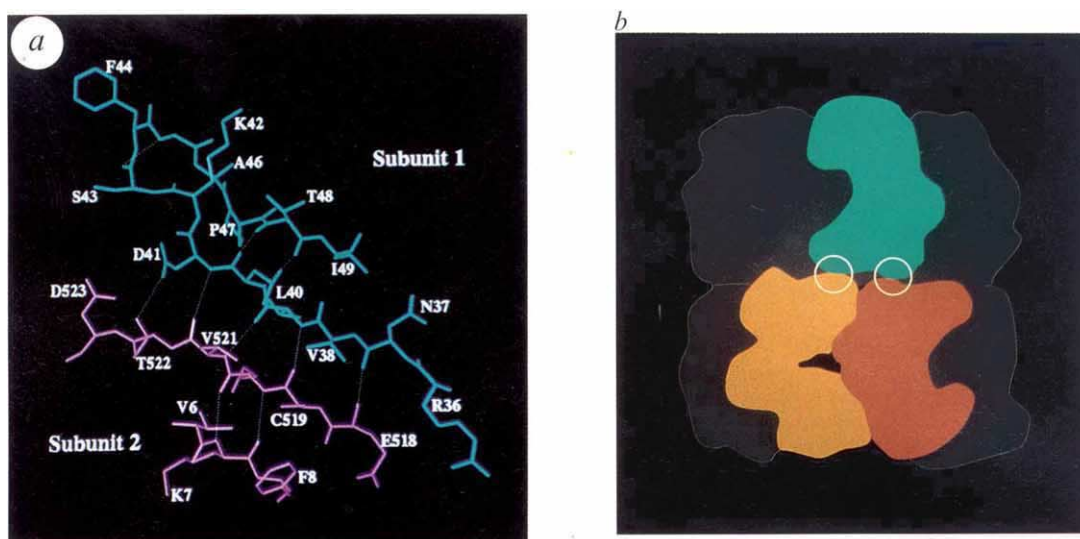


FIG. 4 Intersubunit contacts. *a*, Parallel  $\beta$ -strand interaction between the segment which immediately precedes the disordered C-terminal 24 residues of the 'right' subunit, 518–522 (pink) with an antiparallel loop, 37–49 (blue), which projects from the neighbouring subunit. View is from inside the central channel. This near C-terminal fragment is supported by an antiparallel  $\beta$ -strand interaction with the near N-terminal segment of the same subunit (pink). *b*, Van der Waals surface contacts across the equatorial plane. Top panel highlights in white circles two regions of contact between a subunit in the top ring (blue) and two subunits (orange and red, termed left and right, respectively) in the lower ring. Lower panels show indicated sites of contact as stereo pairs. Residues at the equatorial contact surface are shown with shortened van der Waals radii of 0.9 Å for clarity. These views show nearly exact dyad symmetry. The reason the symmetry is not exact is the following. One of the dyads (not shown here) relating the seven-membered rings corresponds to a lattice symmetry element and, therefore, is crystallographically exact. The remaining six dyads of the assembly are not exact because they are generated by the particle's symmetry which is only roughly 7-fold. These are two different symmetric contact sites because there is an odd number of subunits in the ring. They will be separated by  $360^\circ/7 \div 2 = 25.7^\circ$ .

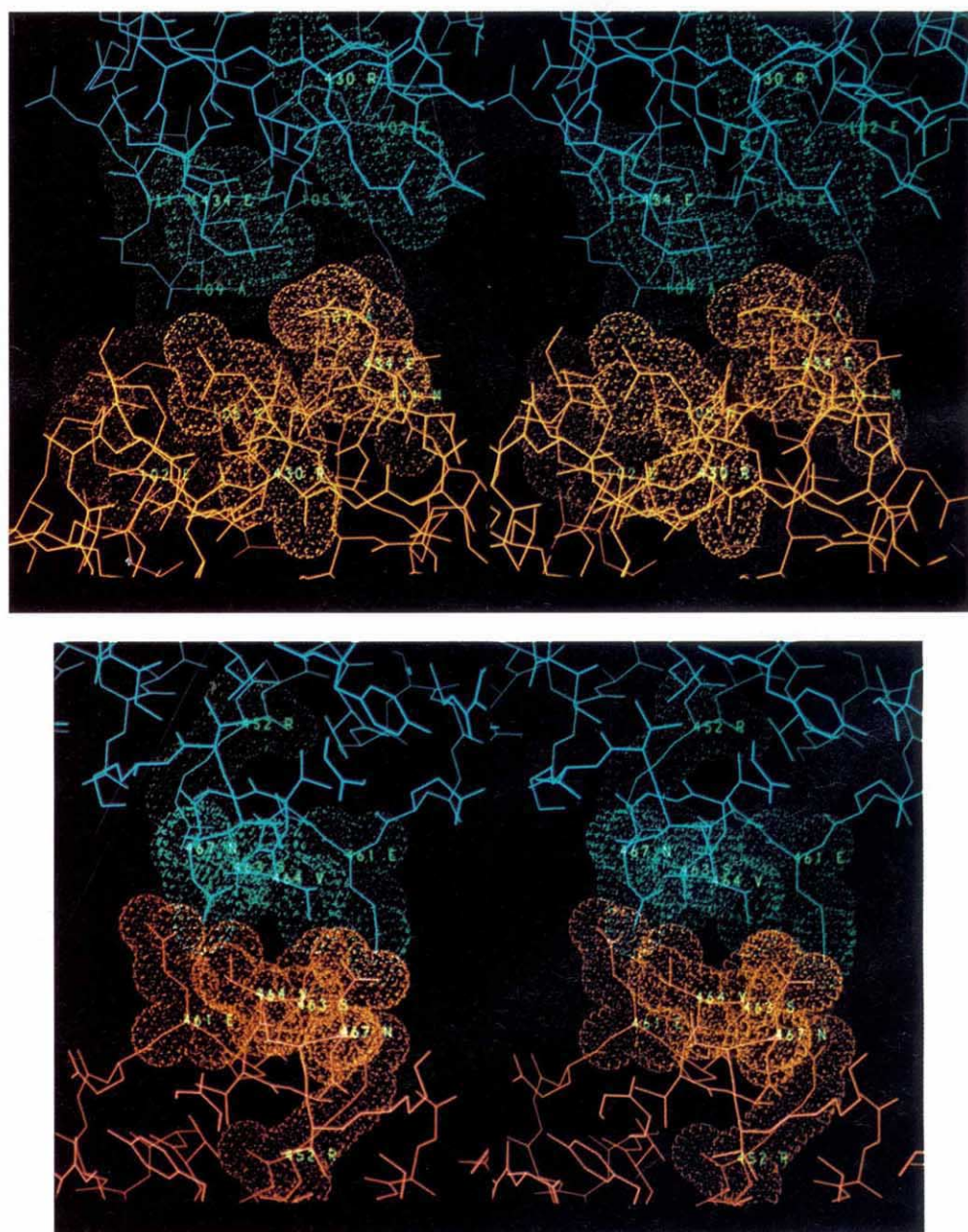


TABLE 2 List of intersubunit contacts

|                                            |             |
|--------------------------------------------|-------------|
| Equatorial loop left/equatorial right      |             |
| Asn 37 scN                                 | Thr 516 scO |
| Asn 37 mcN                                 | Thr 516 mcO |
| Asn 37 mcO                                 | Glu 518 mcN |
| Asp 39 mcN                                 | Cys 519 mcO |
| Asp 39 mcO                                 | Val 521 mcN |
| Asp 41 mcN                                 | Val 521 mcO |
| Asp 41 scO                                 | Thr 522 scO |
| Equatorial left/equatorial right           |             |
| Val 22 sc                                  | Val 6 sc    |
| Arg 36 sc                                  | Met 114 sc  |
| Val 38 sc                                  | Cys 519 sc  |
| Val 39 sc                                  | Met 520 sc  |
| Leu 40 sc                                  | Val 521 sc  |
| Cys 458 mcO                                | Asn 112 scN |
| Intermediate left/apical right             |             |
| Thr 181 mcO                                | Arg 284 mcN |
| Thr 181 mcO                                | Asp 283 scO |
| Met 389 sc                                 | Phe 281 sc  |
| Ala 384 mcO                                | Try 360 scO |
| Glu 386 mcO                                | Arg 197 scN |
| Glu 386 scO                                | Arg 285 scN |
| Leu 183 sc                                 | Tyr 360 sc  |
| Equatorial top ring/equatorial bottom ring |             |
| Left site                                  |             |
| Lys 105 scN                                | Ala 109 mcO |
| Ala 109 sc                                 | Ala 109 sc  |
| Glu 434 scO                                | Glu 434 scO |
| Right site                                 |             |
| Glu 461 scO                                | Arg 452 scN |
| Ser 463 scO                                | Ser 463 scO |
| Val 464 sc                                 | Val 464 sc  |

'sc' and 'mc', Side chain and main chain, respectively. Interactions between polar side chains represent hydrogen bonds between atoms situated 2.6–3.4 Å apart at reasonable angles; interactions between nonpolar side chains represent van der Waals contact distances less than 3.6 Å.

through both rings) is about 146 Å. Were the channel a smooth cylinder, its volume would be  $\sim 250,000 \text{ Å}^3$ . If the partial specific volume of a protein is  $\sim 0.73 \text{ cm}^3 \text{ g}^{-1}$  or  $1.7 \text{ Å}^3$  per unit  $M_r$ , then there is space enough for a polypeptide of  $M_r$  147K provided that it is fitted into the entire channel with perfect close packing. A native folded protein efficiently packed into a crystal lattice requires about  $2.4 \text{ Å}^3$  per unit  $M_r$ , which means the entire channel could accept a folded elliptical protein of  $M_r$   $\sim 100$ K provided its minor width was less than 47 Å. If the volume occupied by a compressed 'unfolded' protein (such as a molten globule) is about double that of its native counterpart, then the central channel could accept 50–60K. However, the process of GroEL-assisted folding presumably requires facile access and egress of the polypeptide substrate, implying a more relaxed rather than highly compressed unfolded state and a smaller value for the mass of protein to be accommodated functionally within the channel would result.

### The side windows

As noted above, each ring is penetrated by seven elliptical windows (portals) that connect the outside to the bulges in the central channel. Unlike the central cavity the windows are not uniform in width, being  $36 \text{ Å} \times 13 \text{ Å}$  at the widest and  $20 \text{ Å} \times 10 \text{ Å}$  at the narrowest. The floor of the portal is the top of the equatorial domain. The oblique left wall and part of the 'roof' is formed by the intermediate domain as it stretches diagonally across to contact the apical domain of its right hand neighbour (Fig. 1c). The ceiling and right wall are composed mainly of the right-hand neighbouring apical domains. Tucked

into the lower left corner toward the channel surface is the ATP-binding site (Fig. 3d of ref. 38 and D.C.B., J. Wang, Z.O., A.L.H. and P.B.S., unpublished results). This portal clearly provides alternative access and egress routes for solution components (such as nucleotides) to the central channel, as well as a possible extension of the peptide binding surface. It may also reduce steric restraint to potential tertiary and quaternary structural adjustments in the allosteric communication between the equatorial and apical domains.

### Discussion

The evidence so far indicates that the main function of GroEL is to provide an interactive surface that offers stable alternative interactions for unfolded and partially folded intermediates that compete favourably against misfolded conformations and/or aggregated states that could not otherwise be reversed on a time-scale suitable for physiologically effective folding. Unlike these so-called 'off-pathway' trajectories, the interactions of the substrate polypeptide with GroEL can be easily disrupted with the aid of ATP and GroES to permit further attempts at the correct folding pathway. If the activation barrier to folding is considered the disruption of incipiently misfolded and/or aggregated states, then the binding of polypeptide to GroEL can be considered part of an enzymatic process whose transition state is kinetically preferable to the off-pathway intermediates in that it provides a rapid potentially productive mode of disengagement.

It is difficult at this stage to draw structure/function conclusions as to how the binding and disengagement process works. This is partly due to the inadequacy of structural information about the complexes formed by GroEL in the reaction pathway, but an equally serious problem is that the course of biochemical events is not well understood. For example, electron microscopy studies suggested a central cavity that might be large enough to allow refolding attempts to occur repetitively within the channel of the same GroEL assembly, as implied by the 'Anfinsen cage' metaphor<sup>46</sup>. Recent biochemical studies, by contrast, indicate that the unfolded/partially folded peptide is completely released to the solvent and re-enters the channel of another GroEL molecule<sup>33,47</sup>. Clearly this implies a different binding and release mechanism. The nature of GroES interactions remains perplexing: for example, does GroES bind to the same ring as polypeptide or to the opposite ring or, as recently suggested<sup>31–33</sup>, to both rings.

Despite these uncertainties, it seems clear from our initial structural and mutational studies that a large portion of the functional surfaces are concentrated on the surface of the inner channel and its invaginations (see Fig. 3 in ref. 38). These include the mutationally defined binding sites on the apical domain for both unfolded substrate polypeptides and GroES, and the ATP-binding pocket on the top surface of the equatorial domain. Although the direct overlap of sites involved in both polypeptide and GroES binding provides a stereochemical basis for mechanisms that couple GroES binding and polypeptide release, the site in the equatorial domain implicated in ATP binding and hydrolysis is separated in the primary structure from the polypeptide-binding segments of the apical domain. Figure 3b, d shows, however, that in the three-dimensional structure, the top of the equatorial domain which bears the ATP binding pocket is close to the undersurface of the apical domain (for example, the carboxylate of Asp 83 is 3.7 Å from the ammonium of Lys 327), suggesting that modest allosteric adjustments transmitted either through the intermediate domain (as noted above) and/or directly through the neighbouring surfaces could couple the presence of the appropriate nucleotide to the process of GroES-assisted polypeptide binding and release.

It remains unclear whether the inside surface of one ring is functionally contiguous with that of the other. Even though they cannot be visualized crystallographically, the C termini of the subunits, totalling 20K per ring, are located within the



central channel (see above). Whether they collectively inhibit the passage of polypeptide between rings is unknown. Deletion of these residues, interestingly, is without apparent effect on GroEL function<sup>48,52</sup>, but this does not exclude the existence of a barrier during normal action. Indeed, it is uncertain that access to the cavity of both rings is even necessary because biochemical studies indicate that polypeptide can be bound by a single ring (ref. 49 and J. Weissman, unpublished data).

Received 30 August; accepted 8 September 1994.

1. Ellis, R. J. & van der Vies, S. M. *Rev. Biochem.* **60**, 321–347 (1991).
2. Gething, M.-J. & Sambrook, J. *Nature* **355**, 33–45 (1992).
3. Hendrick, J. P. & Hartl, F. U. *Rev. Biochem.* **62**, 349–384 (1993).
4. Horwich, A. L. & Willison, K. R. *Phil. Trans. R. Soc., Lond.* **B339**, 313–326 (1993).
5. Hendrix, R. W. *J. molec. Biol.* **129**, 375–392 (1979).
6. McMullen, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **7**, 4414–4423 (1987).
7. Pushkin, A. V. *et al. Biochim. biophys. Acta* **704**, 379–384 (1982).
8. Trent, J. D., Nimmegern, E., Wall, J. S., Hartl, F. U. & Horwich, A. L. *Nature* **354**, 490–493 (1991).
9. Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. H. & Cowan, N. J. *Cell* **69**, 1043–1050 (1992).
10. Fayet, O., Ziegelhoffer, T. & Georgopoulos, C. *J. Bact.* **171**, 1379–1385 (1989).
11. Cheng, M. Y. *et al. Nature* **337**, 620–625 (1989).
12. Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. & Furtak, K. *Cell* **74**, 909–917 (1993).
13. Ursic, D. & Culbertson, M. R. *Molec. cell. Biol.* **11**, 2629–2640 (1991).
14. Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. *Nature* **342**, 884–889 (1989).
15. Martin, J. *et al. Nature* **352**, 36–42 (1991).
16. Mendoza, J. A., Lorimer, G. H. & Horowitz, P. M. *J. biol. Chem.* **266**, 16973–16976 (1991).
17. Bochkareva, E. S., Lissin, N. M., Flynn, G. C., Rothman, J. E. & Girschovich, A. S. *J. biol. Chem.* **267**, 6796–6800 (1992).
18. Landry, S. J., Jordan, R., McMacken, R. & Gierasch, L. M. *Nature* **355**, 455–457 (1992).
19. Zahn, R., Spitzfaden, C., Ottiger, M., Wuthrich, K. & Pluckthun, A. *Nature* **368**, 261–265 (1994).
20. Okazaki, A., Ikura, T., Nikaïdo, K. & Kuwajima, K. *Nature struct. Biol.* **1**, 439–446 (1994).
21. Hayer-Hartl, M. K., Ewbanks, J. J., Creighton, T. E. & Hartl, F. U. *EMBO J.* **13**, 3192–3202 (1994).
22. Langer, T., Pfeifer, G., Martin, J., Baumeister, W. & Hartl, F. U. *EMBO J.* **11**, 4757–4765 (1992).
23. Braig, K., Simon, M., Furuya, F., Hainfeld, J. F. & Horwich, A. L. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3978–3982 (1993).
24. Saibil, H. R. *et al. Curr. Biol.* **3**, 265–273 (1993).
25. Ishii, N., Taguchi, H., Sasabe, H. & Yoshida, M. *J. molec. Biol.* **236**, 691–696 (1994).
26. Gray, T. E. & Fersht, A. R. *FEBS Lett.* **292**, 254–258 (1992).
27. Jackson, G. S. *et al. Biochemistry* **32**, 2554–2563 (1993).
28. Todd, M., Viitanen, P. V. & Lorimer, G. H. *Biochemistry* **32**, 8560–8567 (1993).

This report and the companion mutational study<sup>38</sup> are just a start in dissecting the molecular mechanism of chaperonin-assisted protein folding. Once the molecular events of the pathway are established, a detailed chemical understanding will require structural studies depicting the distribution of bound unfolded polypeptide, the stereochemical consequences of ATP binding and hydrolysis, and the interplay of GroES with GroEL. □

29. Martin, J., Mayhew, M., Langer, T. & Hartl, F. U. *Nature* **366**, 228–233 (1993).
30. Schmidt, M., Buchner, J., Todd, M. J., Lorimer, G. H. & Viitanen, P. V. *J. biol. Chem.* **269**, 10304–10311 (1994).
31. Azem, A., Kessel, M. & Goloubinoff, P. *Science* **265**, 653–656 (1994).
32. Schmidt, M. *et al. Science* **265**, 656–659 (1994).
33. Todd, M., Viitanen, P. V. & Lorimer, G. H. *Science* **265**, 659–666 (1994).
34. Fitzgerald, P. M. D. *J. appl. Crystallogr.* **21**, 273–278 (1988).
35. Otwinowski, Z. in *Proc. CCP4 Study Weekend, 25–26 Jan 1991, Isomorphous Replacement and Anomalous Scattering* (eds Wolf, W., Evans, P. R. & Leslie, A. G. W.) 80–86 (SERC Daresbury Laboratory, UK, 1991).
36. Jones, T. A. *Proc. CCP4 Study Weekend, Molecular Replacement* (eds Dodson, E. J., Glover, S. & Wolf, W. 91–105 (SERC Daresbury Laboratory, UK, 1992).
37. Jones, T. A. *et al. Acta crystallogr.* **A47**, 110–119 (1991).
38. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. *Nature* **371**, 614–619 (1994).
39. Brunger, A. T. *XPLOR Version 3.1 Manual* (1993).
40. Read, R. *Acta crystallogr.* **A42**, 140–149 (1986).
41. Luthy, R., Bowie, J. U. & Eisenberg, D. *Nature* **356**, 83–85 (1992).
42. Lee, B. K. & Richards, F. M. *J. molec. Biol.* **55**, 379–400 (1971).
43. Janin, J. & Chothia, C. *J. biol. Chem.* **265**, 16027–16030 (1990).
44. Janin, J., Miller, S. & Chothia, C. *J. molec. Biol.* **204**, 155–164 (1988).
45. Miller, S. *Prot. Engng* **3**, 77–83 (1989).
46. Ellis, R. J. *Nature* **366**, 213–214 (1993).
47. Weissman, J. S., Kashi, Y., Fenton, W. A. & Horwich, A. L. *Cell* **78**, 693–702 (1994).
48. McLennan, N. F., Girschovich, A. S., Lissin, N. M., Charters, Y. & Masters, M. *Molec. Microbiol.* **7**, 49–58 (1993).
49. Viitanen, P. V. *et al. J. biol. Chem.* **267**, 695–698 (1992).
50. Otwinowski, Z. in *Proc. CCP4 Study Weekend, 29–30 Jan 1991, Data Collection and Processing* (eds Sawyer, L., Isaacs, N. & Burley, S.) 56–62 (SERC Daresbury Laboratory, UK, 1993).
51. Svensson, L. A., Surin, B. P., Dixon, N. E. & Spangfort, M. D. *J. molec. Biol.* **235**, 47–52 (1994).
52. Burnett, B., Horwich, A. L. & Low, K. B. *J. Bact.* (in the press).

ACKNOWLEDGEMENTS. Z.O. and K.B. made equal contributions to this work. We thank J. Wang (Yale) for expert help in implementing the program RAVE, which greatly enhanced portions of the electron-density map; A. Brunger (Yale) for providing access to and help with computation; Y. Kim (Yale) for help with graphics; and S. Ealick (CHESS) and K. Wilson (EMBL Hamburg) for access to their synchrotron facilities. Work was supported in part by grants to P.B.S. and A.L.H. from the NIH. R.H. was a Fellow of the National Cancer Centre.

## LETTERS TO NATURE

# Extremely strong carbon monoxide emission from the Cloverleaf quasar at a redshift of 2.5

Richard Barvainis\*, Linda Tacconi†, Robert Antonucci‡, Danielle Alloin§ & Paul Coleman||

\* MIT Haystack Observatory, Westford, Massachusetts 01886, USA

† Max-Planck-Institut für extraterrestrische Physik, D-85748 Garching, Germany

‡ Physics Department, University of California, Santa Barbara, California 93106, USA

§ URA173 CNRS, Observatoire de Paris, F-92195 Meudon, France

|| Kapteyn Astronomical Institute, Postbus 800, 9700 AV Groningen, The Netherlands

**GALAXIES at high redshift are very faint and difficult to study at optical and near-infrared wavelengths, but detection of far-infrared emission<sup>1</sup> and molecular gas<sup>2,3</sup> in a galaxy at redshift  $z \approx 2.3$  has suggested that their early evolution may be investigated by these means instead. The host galaxies of quasars are promising candidates for these observations, particularly as quasars might be triggered by interactions and mergers between galaxies<sup>4,5</sup> which**

**result in dust- and gas-rich systems. The Cloverleaf, a gravitationally lensed quasar, has far-infrared/submillimetre emission indicating a substantial dust content<sup>6</sup>, and therefore potentially a large amount of gas. Here we report the detection of carbon monoxide emission from the Cloverleaf, which we interpret as indicating a mass of molecular gas that is comparable to the total dynamical mass of the host galaxy, and which is consistent with the total baryonic content of a present-day luminous galaxy. This suggests that, although some processing of gas through stars has taken place in the Cloverleaf at a lookback time of 85% of the current age of the Universe, much of the future stellar content has yet to be formed.**

We used the Institut de Radio Astronomie Millimétrique (IRAM) Interferometer on the Plateau de Bure, France, and the 30-m telescope on Pico Veleta, Spain, for the observations. For the interferometer the dates of observation were 26 and 28 March 1994, the integration time was 7.5 h, and the weather was good. A compact configuration of the four available antennas was chosen to reduce phase noise, and the resulting synthesized beam was  $8.6 \times 7.1$  arcsec, in position angle  $22^\circ$ . The spectrum was centred at 97.134 GHz (lower sideband), corresponding to a redshift of 2.560 for the CO(3–2) line at rest frequency 345.796 GHz. This redshift was chosen as a compromise between the optical H $\alpha$  and [O III] emission lines (H $\alpha$  redshift from P. Barthel, personal communication). The quasars 3C273 and 0923+392 were observed for bandpass calibration, whereas phases and amplitudes were calibrated using the quasar 1413+135 (1.7 Jy at 97 GHz). The interferometer spectrum is

shown in Fig. 1a, where it can be seen that the line is shifted by about  $-125 \text{ km s}^{-1}$  relative to the tuning frequency yielding a line redshift of  $2.5585 \pm 0.0001$  (based on a gaussian fit to the line profile). Maps of individual spectral channels are shown in Fig. 2; no correlation was found between source position and velocity in the channel maps.

The single-dish observations were performed in clear weather on 28 June 1994, with the spectrum centred at 97.172 GHz. System temperatures were in the range 250–300 K, the back-end was a  $512 \times 1 \text{ MHz}$  filter bank, and the integration time was 4.3 h. The subreflector was nutated at 0.5 Hz to provide flat baselines. The CO(3–2) spectrum is shown in Fig. 1b. The derived redshift is  $2.5580 \pm 0.0002$ , which is not significantly different from the interferometer value. The weighted mean redshift is  $2.5584 \pm 0.0001$ . During this observing run we also detected the CO(4–3) and CO(7–6) lines from the Cloverleaf, and the neutral carbon fine-structure  $[\text{C I}](^3\text{P}_1-^3\text{P}_0)$ , eliminating the remote possibility that the signal is from the lensing galaxy rather than the quasar. These results will be presented elsewhere (R.B. *et al.*, manuscript in preparation).

The Cloverleaf derives its name from the optical image, which is split into four spots within  $\sim 1 \text{ arcsec}$  as a result of gravitational lensing<sup>7</sup>. The lensing object has not been detected. The Cloverleaf is also a broad absorption line quasar, and is the only one known to be lensed. The CO and submillimetre properties are similar to those of the  $z=2.3$  superluminous galaxy IRAS F10214+4724, a dusty, molecule-rich object which might also be gravitationally lensed<sup>8</sup>. We are confident that the submillimetre emission from the Cloverleaf is from dust, as the measured cutoff slope of  $2.95 \pm 0.69$  is very steep for self-absorbed synchrotron radiation<sup>6</sup> and the source is radio-quiet. Because the gravitational amplification parameters of the Cloverleaf are not known, calculation of the intrinsic luminosities is uncertain. A simple model<sup>9</sup> assuming a single elliptical galaxy for the lens produces a total amplification of 7.6 for the optical source, summing all four images (this number should be considered representative only). The roughly 1-arcsec Einstein ring diameter is  $3.8h^{-1} \text{ kpc}$  ( $h=H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ), possibly of the order of the size of the dust/CO region, so the dust/CO amplification could be less than in the optical.

In Table 1 we list a number of properties of both the Cloverleaf and IRAS F10214+4724. The derived luminosities and masses must be divided by the unknown amplification factor  $A_G$ . We see that for modest values of  $A_G$  the luminosities and molecular gas content of the objects are comparable. The ratio of CO to (rest) submillimetre luminosities in the Cloverleaf,  $L_{\text{CO}}/(\nu L_{\nu})_{350\mu}$ , is 0.4, which is similar to that found in luminous infrared galaxies and starburst galaxies. For F10214+4724 this

TABLE 1 Properties of the Cloverleaf quasar and F10214+4724

|                                      | Cloverleaf | F10214+4724 | Multiplier                       | Units                             |
|--------------------------------------|------------|-------------|----------------------------------|-----------------------------------|
| $D_L$                                | 10.0       | 8.8         | $\times h^{-1}$                  | Gpc                               |
| $z_{\text{CO}}$                      | 2.5585     | 2.2855      |                                  |                                   |
| $\nu_{\text{CO}}(\text{obs})$        | 97.175     | 105.249     |                                  | GHz                               |
| $S_{\text{CO}}(\text{peak})$         | 23         | 16          |                                  | mJy                               |
| $\Delta\nu_{\text{CO}}(\text{FWHM})$ | 326        | 230         |                                  | $\text{km s}^{-1}$                |
| $S_{\text{CO}}\Delta\nu$             | 8.1        | 4.1         |                                  | Jy $\text{km s}^{-1}$             |
| $L_{\text{CO}}$                      | 8.2        | 3.5         | $\times 10^7 A_G^{-1} h^{-2}$    | $L_{\odot}$                       |
| $L_{\text{CO}}$                      | 6.1        | 2.6         | $\times 10^{10} A_G^{-1} h^{-2}$ | $\text{K km s}^{-1} \text{ pc}^2$ |
| $(\nu L_{\nu})_{350\mu}$             | 13.7       | 6.2         | $\times 10^{10} A_G^{-1} h^{-2}$ | $L_{\odot}$                       |
| $M(\text{H}_2)$                      | 2.4        | 1.0         | $\times 10^{11} A_G^{-1} h^{-2}$ | $M_{\odot}$                       |

Symbols used;  $h=H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ;  $A_G$  is the (unknown) gravitational amplification factor for the CO and submillimetre fluxes,  $D_L$  is the luminosity distance (assuming  $q_0=0.5$ );  $\nu_{\text{CO}}(\text{obs})$ ,  $S_{\text{CO}}(\text{peak})$  and  $\Delta\nu_{\text{CO}}(\text{FWHM})$  are the observed frequency, peak flux density, and velocity full-width to half maximum of the line, respectively, as derived from a gaussian fit to the line profile;  $L_{\text{CO}}$  and  $L_{\text{CO}}$  are the CO luminosities;  $M(\text{H}_2)$  is the derived mass of molecular hydrogen. All values apply to the CO(3–2) line as observed by the Plateau de Bure interferometer. Errors on CO line flux and quantities derived from it are  $\sim \pm 15\%$ . Data for F10214+4724 from refs 3 and 10. See Solomon *et al.*<sup>3</sup> for the discussion of the different ways of expressing CO luminosities, and for discussion of  $M(\text{H}_2)$  as derived from CO measurements. We follow their approach and assume  $M(\text{H}_2)=4L_{\text{CO}}$  (while acknowledging that the conversion factor is somewhat uncertain). The  $350 \mu\text{m}$  (rest) luminosity is based on a 1.25-mm measurement made by us at the IRAM 30-m telescope; it will be presented elsewhere (R.B. *et al.*, manuscript in preparation).

ratio is also 0.4, for Arp220 it is 0.3 and for M82 it is 0.6 (see Table 3 in ref. 10 for a list of these, and other, galaxies).

Mapping high-redshift quasar hosts (and high-redshift galaxies generally) with CO locates quantities of baryonic matter which are significant on galactic scales. This is in contrast to optical studies, in which the diffuse light just establishes the presence of a small mass in upper-main-sequence stars. The CO maps of the Cloverleaf shown in Fig. 2 are consistent with a source that is point-like relative to the beamsize of 8 arcsec, given the signal-to-noise ratio obtained. The limit on the size for the summed channel maps is  $\theta \leq 3.3 \text{ arcsec}$ . The dynamical mass can be expressed as

$$M_{\text{dyn}} = 2 \times 10^5 \Delta\nu_{\text{HWZI}} \theta h^{-1} / \sin(i) N_{\odot}$$

where  $\Delta\nu_{\text{HWZI}}$  is the velocity half-width at zero intensity of the CO line in  $\text{km s}^{-1}$ ,  $\theta$  is the source angular diameter in arcsec,  $i$  is the inclination angle of the system and  $M_{\odot}$  is the solar mass. For  $\Delta\nu_{\text{HWZI}}=200 \text{ km s}^{-1}$  and  $\theta < 3.3 \text{ arcsec}$ ,  $M_{\text{dyn}} < 2.6 \times 10^{10} h^{-1} / \sin(i) M_{\odot}$ . Comparing this to the derived molecular mass from Table 1,  $M(\text{H}_2)=2.4 \times 10^{11} A_G^{-1} h^{-2}$ , we see that most of the mass in the Cloverleaf is in molecular gas as opposed to stars, unless the gravitational amplification of CO is quite large,

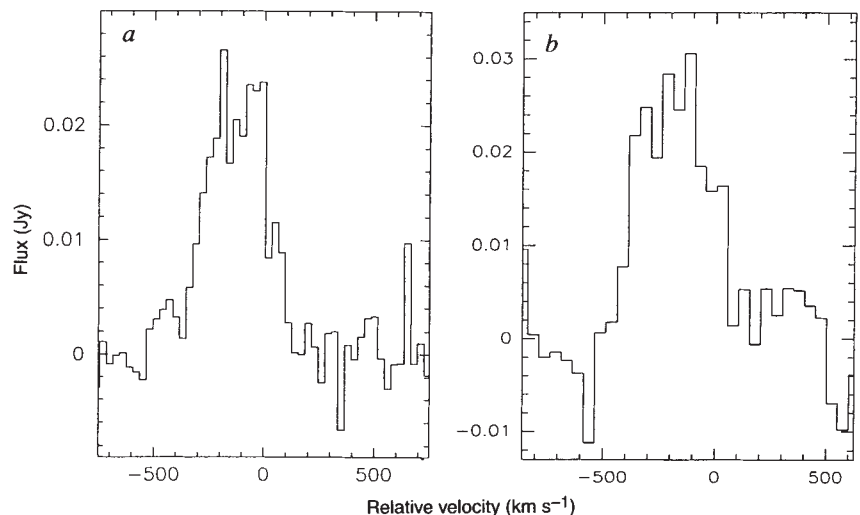


FIG. 1 a, CO  $J=3 \rightarrow 2$  spectrum obtained with the IRAM Plateau de Bure interferometer, smoothed to a resolution of  $30 \text{ km s}^{-1}$ . The velocity scale represents the offset relative to the central observed frequency of 97.134 GHz, in the heliocentric frame. The channel r.m.s. noise is 3 mJy. b, Spectrum taken with the IRAM 30-m telescope, with velocity scale adjusted so as to be relative to 94.134 GHz as in a, and smoothed in this case to  $50 \text{ km s}^{-1}$ . Note slightly different flux density scales. The r.m.s. noise is 5 mJy. The gaussian-fitted peak flux density is  $29 \pm 4 \text{ mJy}$ , compared to  $23 \pm 3 \text{ mJy}$  for the interferometer measurement, and the line width is  $359 \pm 36 \text{ km s}^{-1}$  compared to  $326 \pm 25 \text{ km s}^{-1}$ .



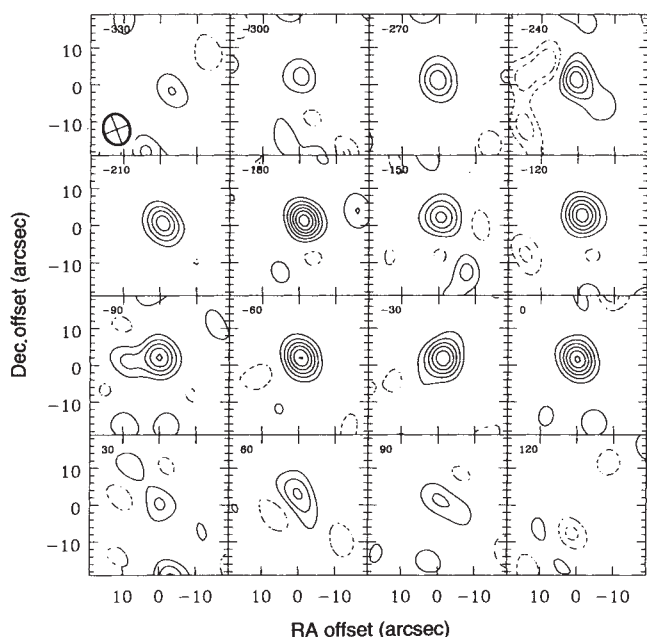


FIG. 2 CO  $J=3 \rightarrow 2$  channel maps. The contours are spaced by 4 mJy. Negative contours are dashed. Relative channel velocities are given at upper left of each frame. Offsets in right ascension (RA) and declination (dec.) are with respect to the position of the optical quasar. The beam size is shown in the upper left-hand frame as a crossed ellipse. Similar maps of the upper spectral sideband (100.1 GHz) show only noise.

or the inclination is quite small. This suggests that we are observing the galaxy in a primitive state, perhaps after an early wave of high-mass star formation has enriched the interstellar medium.

The CO redshift of the Cloverleaf is the only really secure systemic redshift now known for a distant luminous quasar. Systemic redshifts are needed in contexts such as using the "proximity effect" to derive the ambient intergalactic ionizing radiation density<sup>11</sup>, and distinguishing a true intergalactic medium from absorption-line clouds<sup>12</sup>. Our measurement supports the conclusion that the low-ionization broad emission lines and the narrow emission lines are at close to the systemic value, whereas the high-ionization broad lines can have a significantly smaller redshift (see, for example, refs 11, 13). The mean CO redshift is 2.5584, whereas the catalogue value (from the high-ionization lines), is 2.546, a blueshift of  $\approx 1,000 \text{ km s}^{-1}$ . The narrow [O III] 5,007 Å line is at  $z \approx 2.554$ , estimating from the plot in ref. 14. The broad H $\alpha$  line redshift is at 2.554 (from the plot in ref. 14) or 2.562 (P. Barthel, private communication).

Carbon monoxide has been detected previously in several low-redshift quasars<sup>15-18</sup>, and in a number of Seyfert galaxies. These observations have been used to help support two important conclusions. First, the correlation between the CO and far-infrared luminosities seen in luminous IRAS galaxies seems to hold as well for quasars, adding a significant new argument to the list of reasons for believing that the far-infrared emission in radio-quiet quasars and normal (non-blazar) radio-loud quasars is strongly dominated by dust<sup>17,19</sup>. Second, the fact that the CO is so strong in infrared-bright quasars means that large amounts of molecular gas must exist in their host galaxies, comparable to the amounts found in ultraluminous infrared galaxies. This, together with the similar bolometric luminosities and local space densities of quasars and powerful infrared galaxies, and similar

evidence for galactic mergers, has supported the idea that they are really the same type of object but in different evolutionary states, or perhaps seen from different orientations<sup>4,18,20,21</sup>.

These comparisons between quasars and luminous infrared galaxies have been made between objects in the relatively local Universe ( $z < 0.5$ ). With IRAS F10214+4724 at  $z = 2.3$ , and now the Cloverleaf at  $z = 2.5$ , the connection can be made even at the highest luminosities in the early Universe. Comparing data currently available, these objects are extremely similar except in the optical region. There the Cloverleaf is a classical quasar, with both broad and narrow emission lines, whereas F10214+4724 has narrow emission lines similar to a Seyfert 2 nucleus<sup>1,8</sup>. F10214+4724 may be a young, buried quasar, still wrapped in its formative cocoon of dust, or a "quasar 2," the high-luminosity equivalent of the Seyfert 2s. In the latter case it would look much like the Cloverleaf optically if seen from the right orientation, according to unified models for active galactic nuclei<sup>22</sup>. Its high optical polarization suggests that this possibility is the correct one.

The amount of dust and gas present in the Cloverleaf may have been enhanced by a galactic merger or mergers, as has been suggested by many authors for luminous infrared galaxies. At  $z = 2.5$ , the Cloverleaf is probably too far away for us to see direct evidence of a merger, or of companion galaxies, at optical and near-infrared wavelengths. But lower-luminosity quasars certainly do show such signatures, leading to the suggestion that gas dumped into a quasar host during a merger helps both to fuel the quasar and to hide it temporarily. In this scheme the Cloverleaf could be considered a recently emerged quasar, where copious dust and molecules are still present but do not cover the quasar from all angles. Eventually this material is either captured by the central black hole, is blown away, or is turned into stars. □

Received 22 June; accepted 17 August 1994.

- Rowan-Robinson, M. et al. *Nature* **351**, 719–721 (1991).
- Brown, R. L. & Vanden Bout, P. A. *Astrophys. J.* **397**, L19–L22 (1992).
- Solomon, P. M., Downes, D. & Radford, S. J. E. *Astrophys. J.* **398**, L29–L32 (1992).
- Sanders, D. B. et al. *Astrophys. J.* **325**, 74–91 (1988).
- Hutchings, J. B. & Campbell, B. *Nature* **303**, 584–588 (1983).
- Barvainis, R., Antonucci, R. & Coleman, P. *Astrophys. J.* **399**, L19–L22 (1992).
- Magain, P. et al. *Nature* **334**, 325–327 (1988).
- Elston, R. et al. *Astr. J.* **107**, 910–919 (1994).
- Kayser, R. et al. *Astrophys. J.* **364**, 15–22 (1990).
- Downes, D. et al. *Astrophys. J.* **398**, L25–L27 (1992).
- Espey, B. R. *Astrophys. J.* **411**, L59–L62 (1993).
- Jakobsen, P. et al. *Nature* **370**, 35–39 (1994).
- Junkkarinen, V. T. in *Active Galactic Nuclei* (eds Osterbrock, D. E. & Miller, J. S.) 122 (Int. Astr. Un. Symp. No. 134, Kluwer, Dordrecht, 1989).

- Hill, G. J., Thompson, K. L. & Elston, R. *Astrophys. J.* **414**, L1–L4 (1993).
- Sanders, D., Scoville, N. Z. & Soifer, B. T. *Astrophys. J.* **335**, L1–L4 (1988).
- Barvainis, R., Alloin, D. & Antonucci, R. *Astrophys. J.* **337**, L69–L72 (1989).
- Alloin, D., Barvainis, R., Gordon, M. & Antonucci, R. *Astr. Astrophys.* **265**, 429–436 (1992).
- Scoville, N. Z., Padin, S., Sanders, D. B., Soifer, B. T. & Yun, M. S. *Astrophys. J.* **415**, L75–L77 (1993).
- Barvainis, R. in *Testing the AGN Paradigm* (eds Holt, S. S., Neff, S. G. & Urry, C. M.) 129–138 (AIP, New York, 1992).
- Antonucci, R. & Olszewski, E. W. *Astr. J.* **90**, 2203–2206 (1985).
- Barvainis, R. *Nature* **366**, 298–299 (1993).
- Antonucci, R. A. *Rev. Astr. Astrophys.* **31**, 473–521 (1993).

ACKNOWLEDGEMENTS. We thank S. Guilloteau for scheduling our project, and D. Downes for assistance with the data reduction.

# Possible evidence against a massive black hole at the Galactic Centre

**A. Goldwurm\***, **B. Cordier\***, **J. Paul\***, **J. Ballet\***,  
**L. Bouchet†**, **J.-P. Roques†**, **G. Vedrenne†**,  
**P. Mandrou†**, **R. Sunyaev†**, **E. Churazov†**,  
**M. Gilfanov†**, **A. Finogonov†**, **A. Vikhlinin†**,  
**A. Dyachkov†**, **N. Khavenson†** & **V. Kovtunen†**

\* CEA/DSM/DAPNIA/Service d'Astrophysique,  
Centre d'Etudes de Saclay, 91191 Gif-sur-Yvette Cedex, France

† Centre d'Etude Spatiale des Rayonnements,  
9, avenue du Colonel Roche, BP 4346, 31029 Toulouse Cedex,  
France

‡ Space Research Institute, Profsoyuznaya, 84/32, Moscow 117296,  
Russia

THE centre of our Galaxy is known to contain a large condensation of mass<sup>1</sup>, and it has been suggested that a massive black hole (of several million solar masses) is located there. Massive black holes have been proposed to explain active galactic nuclei, and if the Galactic Centre is a less-powerful version of such sources it should radiate X-rays and  $\gamma$ -rays<sup>1</sup>. But although earlier observations<sup>2-6</sup> have shown that the region does emit X-rays and  $\gamma$ -rays, the true centre, corresponding to the object Sagittarius A\*, does not emit strongly at least up to energies of 30 keV (refs 4-6). Whether Sgr A\* emits radiation at higher energies, however, was not resolved. Here we present the results of a deep imaging survey of the Galactic Centre, performed with the Sigma/GRANAT telescope. We determine the locations of the nine hard X-ray sources—six of them being observed for the first time in the spectral band—to an accuracy of about 2 arcmin, but find no source associated with Sgr A\*. The hard X-ray luminosity of Sgr A\* is a factor of  $4 \times 10^7$  less than that expected for a black hole of a million solar masses accreting gas at the maximum stable rate, challenging the idea that there is a black hole at the Galactic Centre.

Since its launch aboard the Russian GRANAT spacecraft on 1 December 1989, the French telescope Sigma<sup>7</sup> has observed twice per year the Galactic Centre region in the 35–1,300 keV band. The Sigma imaging performances are obtained by the combination of a position-sensitive  $\gamma$ -ray camera, composed of a single NaI crystal, and a URA tungsten coded mask placed above the detector plane. The telescope angular resolution is  $\sim 15$  arcmin, but point sources can be located with an accuracy of 30 arcsec to 5 arcmin, depending on the signal-to-noise ratio. Its large field of view, up to  $18.0^\circ \times 16.7^\circ$  at zero sensitivity, and the large amount of observing time available allowed us to map large regions of the sky in hard X-rays and soft  $\gamma$ -rays with unprecedented precision. Our first observations of the Galactic Centre substantiated the results obtained by the two balloon-borne experiments GRIP and EXITE suggesting that IE1740.7–2942 was visible up to 100 keV (refs 8, 9), and clinched its identification as the major source of high-energy emission in the central  $1^\circ$  (refs 10, 11), and even as a possible site of positron production<sup>12, 14</sup>. Sigma also discovered the very hard source GRS1758–158 (ref. 15), and detected flares of hard radiation from several other well known X-ray sources in the region.

We have recently refined the procedures for analysing Sigma images<sup>16</sup> and re-analysed all the data collected by Sigma from the  $\sim 18^\circ \times 17^\circ$  region around the Galactic Centre, between 24 March 1990 and 14 October 1993. (A total of 109 observing periods, for an effective dead-time-corrected integration time of 1,752 hours, have been analysed.) This unprecedented survey of the Galactic Centre region, spanning nearly 4 years, has yielded the most precise images available beyond 35 keV, with a typical  $1\sigma$  flux error in the adopted energy bands of  $< 2-3$  mCrab (1 mCrab corresponds to  $8.0 \times 10^{-12}$  erg cm<sup>-2</sup> s<sup>-1</sup> in the 35–75 keV band and to  $6.9 \times 10^{-12}$  erg cm<sup>-2</sup> s<sup>-1</sup> between 75 and 150 keV). The resulting images in the 35–75 keV and 75–150 keV bands are displayed in Fig. 1. The low-energy image (Fig. 1a) shows a cluster of well-separated sources detected above the  $5\sigma$  level. Their estimated positions and average fluxes are given in Table 1, along with their proposed identifications<sup>17, 21</sup>. Note that the strongest one is the hard X-ray nova which we discovered in Ophiuchus on 25 September 1993 (ref. 17). While flaring it

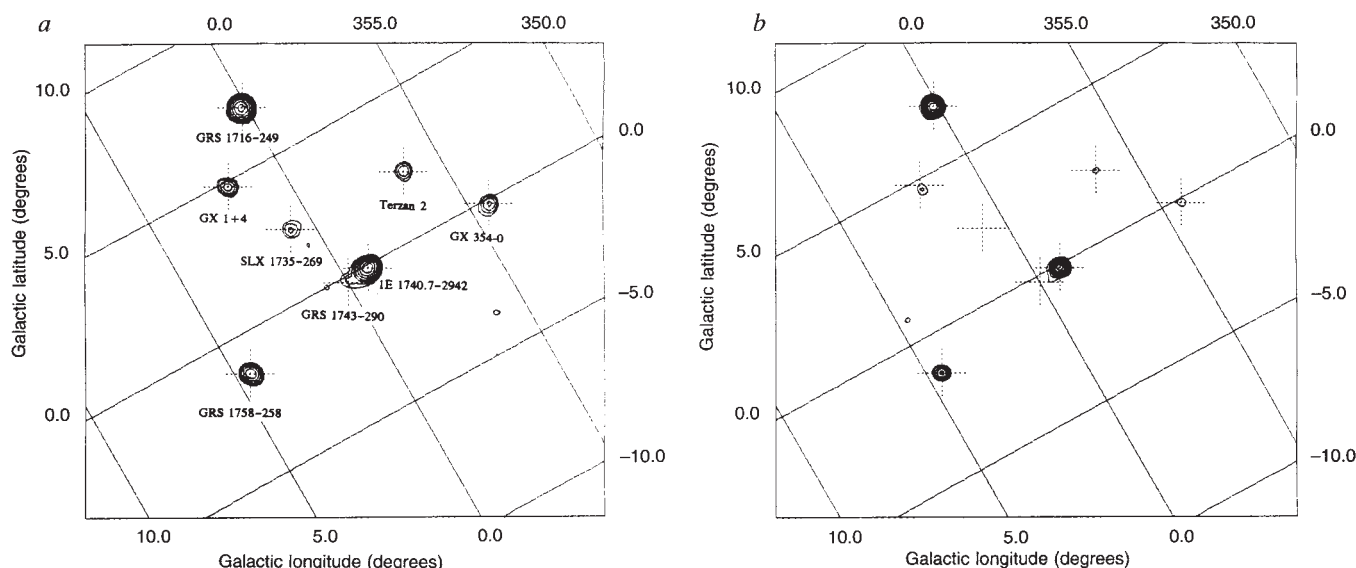


FIG. 1 Images of the  $18.0^\circ \times 16.7^\circ$  sky region around the Galactic Centre as seen by the Sigma/GRANAT telescope after nearly 4 years of operation in the 35–75 keV (a) and the 75–150 keV (b) bands. Images, after standard corrections and deconvolution, have been smoothed with the telescope PSF and summed. Contour levels are in units of local standard deviations ( $\sigma$ ), starting at  $4.5\sigma$  and logarith-

mically spaced by a factor 1.382 ( $4.5\sigma$ ,  $6.2\sigma$ ,  $8.6\sigma$ , ...,  $60\sigma$ ). The sensitivity is not uniform along the image, being maximal around the Galactic nucleus and decreasing regularly to 15% of the maximum at the edges. Crosses indicate positions of the detected sources. For clarity, the position of A1742–294 is not reported (see Fig. 2).



TABLE 1 Sources detected in Galactic Centre survey

| Source<br>Sigma reference | Sigma position<br>(1950 equinox)       |                       |  | Fluxes and $1\sigma$ upper limits<br>(mCrab) |                 |      | Nature               |
|---------------------------|----------------------------------------|-----------------------|--|----------------------------------------------|-----------------|------|----------------------|
|                           | RA (h min s)<br>Dec. ( $^{\circ}$ ' ") | Err. rad.<br>90% c.l. |  | 35–75 keV                                    | 75–150 keV      | HR   |                      |
| GRS1716–249*              | 17 16 34                               | 1.0'                  |  | $139.0 \pm 2.8$                              | $175.8 \pm 3.3$ | 1.3  | Black hole candidate |
| (ref. 17)                 | –24 57 45                              |                       |  |                                              |                 |      |                      |
| Terzan 2                  | 17 24 13                               | 1.8'                  |  | $21.7 \pm 1.9$                               | $12.1 \pm 2.2$  | 0.6  | X-ray burster        |
| (ref. 18)                 | –30 44 37                              |                       |  |                                              |                 |      |                      |
| GX354–0                   | 17 28 47                               | 1.7'                  |  | $25.8 \pm 2.1$                               | $13.5 \pm 2.4$  | 0.5  | X-ray burster        |
| (ref. 19)                 | –33 47 47                              |                       |  |                                              |                 |      |                      |
| GX1+4                     | 17 29 03                               | 1.5'                  |  | $29.2 \pm 2.1$                               | $12.2 \pm 2.4$  | 0.4  | Accreting            |
| (ref. 19)                 | –24 42 18                              |                       |  |                                              |                 |      | X-ray pulsar         |
| SLX1735–269               | 17 35 13                               | 3'                    |  | $14.8 \pm 1.6$                               | $5.2 \pm 1.9$   | 0.4  | Neutron star?        |
|                           | –26 58 38                              |                       |  |                                              |                 |      |                      |
| 1E1740.7–2942*            | 17 40 42                               | 27"                   |  | $65.5 \pm 1.6$                               | $69.9 \pm 1.9$  | 1.1  | Black hole candidate |
| (ref. 11)                 | –29 43 23                              |                       |  |                                              |                 |      |                      |
| A1742–294*                | 17 42 43                               | 5'                    |  | $8.0 \pm 1.6$                                | $4.9 \pm 1.9$   | 0.6  | X-ray burster        |
| (ref. 21)                 | –29 29 06                              |                       |  |                                              |                 |      |                      |
| GRS1743–290               | 17 43 10                               | 4'                    |  | $10.1 \pm 1.6$                               | <1.9            | <0.2 | Neutron star?        |
|                           | –29 02 22                              |                       |  |                                              |                 |      |                      |
| GRS1758–258*              | 17 58 07                               | 40"                   |  | $42.2 \pm 2.2$                               | $40.4 \pm 1.9$  | 1.0  | Black hole candidate |
| (ref. 15)                 | –25 44 30                              |                       |  |                                              |                 |      |                      |

Positions, intensities and hardness ratio (HR) of sources detected by Sigma/GRANAT in the Galactic Centre survey. Positions (right ascension, RA; declination, dec.) with their 90% confidence level (c.l.) error radius (err. rad.) in 2 parameters have been estimated in the present work from the data of Fig. 1a, except for sources labelled with an asterisk, in which cases we retained the more precise position reported in the reference quoted in the first column.

was so bright ( $\sim 1$  Crab) that even averaging over the whole 1990–93 period, its significance remains the highest.

The 35–75 keV emission detected in the central  $1^{\circ}$  radius circle in Fig. 1a, although dominated by 1E1740.7–2942, is not compatible with the telescope point spread function (PSF), suggesting the presence of multiple sources. Inspection of images in different periods allowed us to resolve the residual signal from two other variable point sources. Part of it is due to the emission of the close ( $\sim 30$  arcmin) X-ray burster A1742–294 in autumn 1992 and 1993 (ref. 21). The remaining signal can be associated with a new source, localized east of Sgr A\*, which was brighter in 1991 ( $\sim 20$  mCrab) and probably active also in 1990 and beginning of 1992. By using the images collected in 1991, when

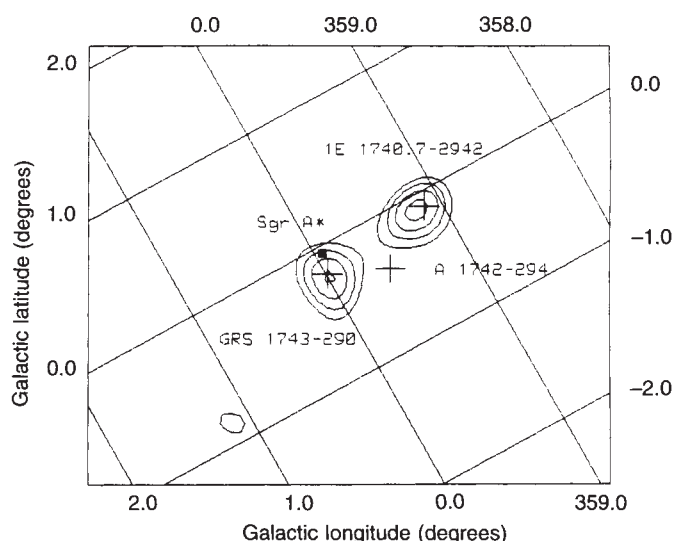


FIG. 2 Image of the sky region around the Galactic nucleus in the energy band 35–75 keV as seen by Sigma during 1991. Contour levels are in units of local standard deviations at  $3.25\sigma$ ,  $4.00\sigma$ ,  $4.75\sigma$  and  $5.5\sigma$ . Crosses indicate the best estimates of the positions of 1E1740.7–2942, A1742–294 and the fitted position of the new source (GRS1743–290). The position of Sgr A\* is shown by a filled square.

1E1740.7–2942 was particularly weak<sup>11</sup> and A1742–294 not visible at all, we could unambiguously localize this excess. As shown in Fig. 2, it appears at  $5.7\sigma$ , with a shape compatible with the PSF, and well separated from 1E1740.7–2942. The estimated position of this new source, GRS1743–290, as derived from this image by fitting the two sources simultaneously, is given in Table 1. The 90% confidence level error box of 4.0 arcmin size (in two parameters) is not significantly influenced by the presence,  $\sim 50$  arcmin away, of 1E1740.7–2942. Sgr A\* is the closest known  $>3$  keV source to GRS1743–290, but the offset ( $\sim 9.4$  arcmin) appears much too large to identify it with the new source. The radio transient detected late in 1990 (ref. 22) is also too far away, and the closest Einstein source<sup>3</sup> (1E1742.7–2902) is an implausible candidate because of its weakness and softness. The fact that no counterpart of GRS1743–290 is visible in the maps obtained by the ART-P telescope aboard GRANAT in 1991 (ref. 6) suggests that this source could be deeply embedded in, or behind, a dense molecular cloud. Using a  $E^{-\alpha}$  power-law spectrum with  $2 < \alpha < 3$  (typical of ART-P sources), column densities in the range  $(1.6\text{--}4) \times 10^{24} \text{ cm}^{-2}$  are needed to reduce the extrapolated 8–20 keV flux below the  $3\sigma$  ART-P sensitivity. Such column density can be provided by the dense part of the Sgr A complex, positioned along the line of sight of the source as indicated by infrared observations<sup>23</sup>. Part of the scattered source flux must actually escape the cloud and contribute to the 7–20 keV diffuse emission of the region, which was recorded by ART-P in 1990 (ref. 24).

Out of the  $\sim 60$  sources known to populate the observed region in the 2–20 keV band at the beginning of the Sigma survey, only 6 have been clearly detected beyond 35 keV in our long exposure. With the exception of the hard X-ray nova in Ophiuchus, the sources listed in Table 1 are probably responsible for the bulk of the hard X-ray emission observed in the region in the past 20 years. On the basis of the hardness ratio—the ratio between the 75–150 keV and the 35–75 keV count rates (see Table 1)—they can be classified in two distinct groups. The hard group, for which the hardness ratio  $\geq 1$ , includes the three brighter sources, all potential black hole candidates, because of the similarity of their spectra with that of Cygnus X-1, the most popular black hole candidate. Indeed, hard spectra going well beyond 100 keV have been considered as spectral signatures of black holes. The basic argument is that the high temperatures required (for

example, the electron temperature of a hot plasma generating hard radiation by Compton collisions) cannot be easily maintained in neutron star systems where strong, low-energy emission from the star surface would cool the plasma by Compton scattering<sup>10</sup>. The soft group (hardness ratio  $\leq 0.6$ ), includes four genuine neutron-star binaries: one accreting X-ray pulsar, and three X-ray bursters. These sources have already been reported to feature sporadic hard X-ray emission, but our analysis shows now that they must actually be considered nearly persistent sources of hard X-rays, systematically fainter and softer than black hole candidates. Using the estimate of the hardness ratio as a criterion to distinguish the nature of hard X-ray emitting compact objects, we propose that the two sources SLX1735–269 and GRS1743–290 are also accreting neutron stars. Note also that the sporadic broad high-energy feature observed on three occasions by Sigma from 1E1740.7–2942 (refs 12–14) and probably related to positron annihilation processes substantiates the hypothesis of the black-hole nature of this source.

An important result of the Sigma deep survey of the Galactic Centre is that the Galactic nucleus is silent above 35 keV. Although in the 75–150 keV band, the 1990–93 average luminosity is definitely  $<2.4 \times 10^{35} \text{ erg s}^{-1}$  ( $2\sigma$  upper limit for a distance of 8.5 kpc), it is more difficult to determine such a compelling limit in the 35–75 keV band, because of the presence of the nearby source GRS1743–290. Although we can reject the hypothesis that Sgr A\* is the counterpart of GRS1743–290 (probability  $<10^{-5}$  for two degrees of freedom), the upper limit of the 35–75 keV luminosity of the Galactic nucleus cannot be set below  $3.5 \times 10^{35} \text{ erg s}^{-1}$ . These limits can be compared to the nearly contemporary data obtained by ART-P during autumn 1990 (ref. 6). The rather flat spectrum ( $\alpha=1.6$ ) of Sgr A\* measured by ART-P (and extrapolated to the adopted Sigma energy bands) implies luminosities respectively 3 and 5.7 times higher than the quoted upper limits in the 35–75 and the 75–150 keV bands. Even considering the uncertainties in the ART-P spectrum, our data imply a break in the Sgr A\* X-ray spectrum at a critical energy, probably between 20 and 40 keV. As a consequence, the total 3–150 keV luminosity of Sgr A\* cannot exceed  $2.5 \times 10^{36} \text{ erg s}^{-1}$ , a value which must be compared to the Eddington luminosity of a black hole of  $10^6$  solar masses that is,  $\sim 10^{44} \text{ erg s}^{-1}$ . In the hypothesis that a massive black hole at the Galactic nucleus accretes matter from the close ( $\sim 0.04 \text{ pc}$ ) hot star cluster IRS156 via powerful stellar winds, the limit on the black hole mass which can be derived from our hard X-ray upper limits, and from the break we predict in the X-ray spectrum, is still rather controversial<sup>25,26</sup>, because it depends strongly on the assumptions made about the model and radiation mechanisms. Our results definitely show that if an accreting massive black hole resides at the Galactic nucleus, it does not efficiently convert to hard X-ray and soft  $\gamma$ -rays the potential energy provided by accretion of the IRS16 star-cluster wind, and it clearly does not behave like a scaled-down active galactic nucleus. In addition, past detection of a variable narrow 511 keV line emission from the Galactic Centre region, although not localized, has often supported the belief that a massive black hole resides in the Galactic nucleus<sup>1</sup>. The lack of 511 keV line detection from the Galactic nucleus itself—a Sigma upper limit of  $2.45 \times 10^{-4} \text{ photons cm}^{-2} \text{ s}^{-1}$  has been obtained for both 1E1740.7–2942 and Sgr A\* (ref. 27)—does not rule out the possible presence of a weak compact variable 511 keV source, but certainly implies that the Galactic nucleus remains underluminous at the electron-positron annihilation energy. □

7. Paul, J. et al. *Adv. Space Res.* **11**, 289–302 (1991).
8. Cook, W. R. et al. *Astrophys. J.* **372**, L75–L78 (1991).
9. Covault, C., Manandhar, R. & Grindlay, J. *Proc. 22nd Int. Cosmic Ray Conf. Dublin Vol. 1* 21–24 (1991).
10. Sunyaev, R. et al. *Astr. Astrophys.* **247**, L29–L32 (1991).
11. Cordier, B. et al. *Astr. Astrophys.* **272**, 277–284 (1993).
12. Bouchet, L. et al. *Astrophys. J.* **383**, L45–L48 (1991).
13. Sunyaev, R. et al. *Astrophys. J.* **383**, L49–L52 (1991).
14. Cordier, B. et al. *Astr. Astrophys.* **275**, L1–L4 (1993).
15. Gilfanov, M. et al. *Astrophys. J.* **418**, 844–849 (1993).
16. Goldwurm, A. in *Proc. of Capri Workshop on Imaging in High Energy Astronomy* (eds Bassani, L. & Di Cocco, G.) (Bologna, Italy, in the press).
17. Gilfanov, M. et al. in *Proc. 4th A. Astr. Maryland Conf. College Park, Maryland, USA* (in the press).
18. Barret, D. et al. *Astrophys. J.* **379**, L21–L24 (1991).
19. Claret, A. et al. *Astrophys. J.* **423**, 436–440 (1994).
20. Laurent, P. et al. *Astr. Astrophys.* **278**, 444–448 (1993).
21. Churazov, E. et al. *Astrophys. J. Suppl. Ser.* **92**, 381–385 (1994).
22. Zhao, J. et al. *Science* **255**, 1538–1543 (1992).
23. Cox, P. & Laureijs, R. in *The Center of the Galaxy* (ed. Morris, M.) 121–128 (IAU Kluwer, Symp. No. 136, Dordrecht, 1989).
24. Markevitch, M., Sunyaev, R. A. & Pavlinsky, M. N. *Nature* **364**, 40–42 (1993).
25. Mastichiadis, A. & Ozerov, L. M. *Astrophys. J.* **426**, 599–603 (1994).
26. Melia, F. *Astrophys. J.* **426**, 577–585 (1994).
27. Malet, I. et al. *Astrophys. J.* (in the press).

ACKNOWLEDGEMENTS. We acknowledge the vital contribution of the CNES Toulouse Space Centre to this work, and thank the staff at the Lavatchine Space Company, the Babakin Space Centre, the Baikonour Space Centre and the Evpatoria Ground Station for their unfailing support.

## An organic solid with wide channels based on hydrogen bonding between macrocycles

D. Venkataraman\*, Stephen Lee†‡, Jinshan Zhang†§ & Jeffrey S. Moore\*‡

\* Departments of Chemistry and Materials Science & Engineering, University of Illinois, Urbana, Illinois 61801, USA

† Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA

RESEARCH on microporous solids has focused largely on inorganic materials such as aluminosilicates (zeolites), aluminophosphates, pillared clays and other layered materials<sup>1,2</sup>. An elusive goal has been the design of new materials with specific properties such as selective adsorption and catalytic activity. It would be very useful if the tools of molecular synthesis could be brought to bear on this problem. Here we report the design, based on a modular approach, and the crystal structure of an organic solid with large-diameter (about 9 Å) extended channels. The channels are formed from planar, rigid macrocyclic building blocks. Onto the outer rim of the macrocycles are attached phenolic groups, which form hexagonally closest-packed two-dimensional hydrogen-bonded networks. Extended channels result from the stacking of these layers in a way that maintains registry between the macrocyclic cavities, and these channels are filled with solvent molecules. This approach potentially offers a simple means to exercise control over pore size and shape in the solid state.

One of the difficulties in applying the methodology of organic chemistry to the design of porous solids is the isotropic nature of intermolecular interactions of organic molecules. This isotropy generally leads to closest packings. Work has been carried out on the use of more directional intermolecular interactions, such as hydrogen bonds, to control solid-state organization<sup>3–7</sup>. Many examples of one-dimensional chains as well as two- and three-dimensional hydrogen-bonded networks are known. Networks have even been designed with the intention of creating large cavities; however, in the solid state the networks tend to self-

Received 1 June; accepted 8 September 1994.

1. Genzel, R. & Townes, C. H. *Astr. Astrophys.* **25**, 377–423 (1987).
2. Gehrels, N. & Tueller, J. *Astrophys. J.* **407**, 597–605 (1993).
3. Watson, M. G., Willingale, R., Grindlay, J. & Hertz, P. *Astrophys. J.* **250**, 142–154 (1981).
4. Skinner, G. K. et al. *Nature* **330**, 544–547 (1987).
5. Hertz, P. & Grindlay, J. E. *Astrophys. J.* **278**, 137–149 (1984).
6. Pavlinsky, M. N., Grebenev, S. A. & Sunyaev, R. *Astrophys. J.* **425**, 110–121 (1994).

‡ Authors to whom correspondence should be addressed.

§ Present address: Motorola Inc., 8000 West Sunrise Boulevard, Fort Lauderdale, Florida 33322, USA.



interpenetrate, filling the voids left in the initial host structure<sup>8-13</sup>. Generally the end result is a dense structure, although a nanoporous coordination network was reported recently by Abrahams *et al.*<sup>14</sup>. Solid-state organic hosts which enclose guests in extended channels are well known, but usually these channels are formed by intermolecular packing voids which happen to lie along a common axis<sup>15,16</sup>. Only a small number of tubular clathrates are known in which the walls of the channel are formed by the interior of a macrocyclic building block<sup>17-19</sup>.

We suggest that a promising approach to molecular-based porous crystals would involve the use of high-symmetry, multi-coordinating macrocycles made of stiff units that resist buckling even in the solid state. Depending on the nature of the macrocycle, even closest-packed motifs of such building blocks could exhibit significant porosity, with the added advantage that closest packing would prevent interpenetration. Hence macrocycle **1** (Fig. 1) with phenolic groups on the exterior was synthesized<sup>20</sup>, with the expectation that the phenolic groups would organize intermolecular hydrogen bonding, leading to a two-dimensional hexagonal network. Based on the known aggregation behaviour<sup>21,22</sup> of the macrocycles, it was postulated that these layers might pack atop one another in such a fashion that extended channel structures would be generated.

Crystals were grown by the slow diffusion of methanol into a solution of **1** in ethanol in a 1-mm sealed capillary tube. The lattice was found to be trigonal with cell dimensions  $a = 20.660(8)$  and  $c = 9.998(7)$  Å. The systematic extinctions, together with  $R_{\text{merge}}$ , showed the space group to be  $P3_1$  (no. 144) or equivalently  $P3_2$  (no. 145). As the molecule itself is a toroid of approximately  $18 \times 18 \times 3.5$  Å, it is clear by comparison to the unit cell dimensions that no more than three molecules can occupy the primitive cell. Furthermore, as there is only one site of symmetry in  $P3_2$ , and this site has 3-fold multiplicity, we concluded that there is only one crystallographically inequivalent molecule per unit cell. The orientation of this single molecule can be unambiguously deduced from the Patterson function (note that as the molecule is of  $D_{6h}$  symmetry and as the molecules lie normal to the  $c$  axis, all bonds of **1** are in the same directions of the Patterson map). Molecule **1** is nearly rigid, and so the only remaining variables needed for an initial solution were the  $x$  and  $y$  coordinates of the centre of the crystallographically inequivalent molecule. They were determined by systematic trial and error. Allowing the molecule to relax from a perfectly planar geometry resulted in an  $R$  factor ( $R_1$ ) of 32.4%. At this point, solvent molecules consisting of 12 ethanol, 6 methanol and 3 water molecules appeared in the Fourier map (Fig. 2). The density observed by the flotation method was close to the

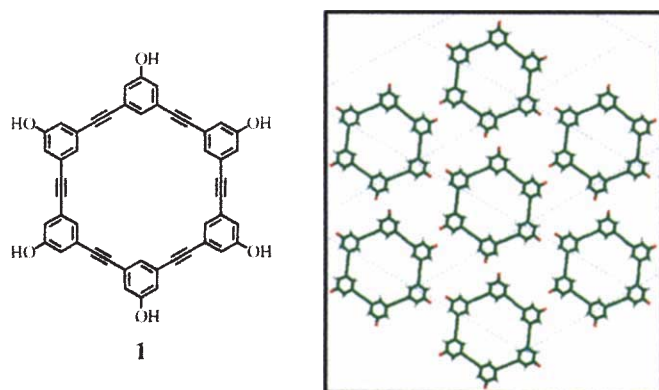


FIG. 1 Left, molecular formula of compound **1**. Right, single layer motif from the crystal structure of **1** parallel to the (001) plane. The macrocycles form a two-dimensional hexagonal closest packing stabilized by hydrogen-bonding interactions. Included solvent molecules have been removed for clarity.

TABLE 1 Crystal data for compound **1** at  $-140^\circ\text{C}$

|                                          |                                                                                                  |
|------------------------------------------|--------------------------------------------------------------------------------------------------|
| Crystal system                           | Trigonal                                                                                         |
| Space group                              | $P3_1(144)$ or $P3_2(145)$                                                                       |
| Cell parameters                          | $a = b = 20.660(9)$ Å<br>$c = 9.998(7)$ Å<br>$\alpha = \beta = 90^\circ$<br>$\gamma = 120^\circ$ |
| Z                                        | 3                                                                                                |
| Volume                                   | $3,695.5(33)$ Å <sup>3</sup>                                                                     |
| Octants used for data collection         | Full sphere                                                                                      |
| No. of reflections refined               | 2,704                                                                                            |
| No. of parameters                        | 130                                                                                              |
| Unweighted agreement factor ( $R_1$ )    | 0.1758 ( $F > 4\sigma$ )                                                                         |
|                                          | 0.1816 (all data)                                                                                |
| Highest peak in the final difference map | 0.67 electrons per Å <sup>3</sup>                                                                |
| Goodness of fit                          | 1.157                                                                                            |

Experimental details: diffractometer, Syntex P2<sub>1</sub> IV; Mo  $\mu$   $\lambda(\text{Mo K}\alpha) = 0.075$  nm<sup>-1</sup>; (Mo  $\text{K}\alpha$ ) =  $0.710734$  Å,  $L_p$  corrected; 8,246 reflections collected; number of reflections  $> 4.0$  ( $\sigma(f)$ ) = 2,391; computing structure solution, SHELXTL PLUS; computing structure refinement, SHELXL-93; reflections were refined based on  $F_o^2$  by full matrix least-squares.

calculated density (calculated,  $1.29$  g cm<sup>-3</sup>; observed,  $1.24$  g cm<sup>-3</sup>). Solvent molecules were at positions compatible with hydrogen bonding to the hexaphenol and with reasonable bond distances and bond angles for ethanol and methanol. Disorder in the solvent atoms was apparent from their thermal parameters, especially for those carbons furthest removed from the hydrogen-bonded oxygen. The disorder observed in the solvent is in sharp contrast to macrocyclic atoms which were well resolved in subsequent electron density maps. Isotropic refinement led to a final  $R_1$  factor of 17.58 for 2,391 reflections ( $F > 4.0\sigma$ ) and 130 parameters (Table 1)<sup>23</sup>. The bond angles and distances are in accord with the chemical structure of the molecule. The Fourier difference map showed a residual electron density of  $0.67$  electrons Å<sup>-3</sup>.

The structure is layered, with each layer consisting of sheets of macrocycles hydrogen-bonded to one another forming a two-dimensional closest-packed structure (Fig. 1). Of the two common types of aromatic packing modes<sup>24</sup>, namely the stack and herringbone patterns, efficient packing in the latter is prevented by the topological constraints of macrocyclization and the drive

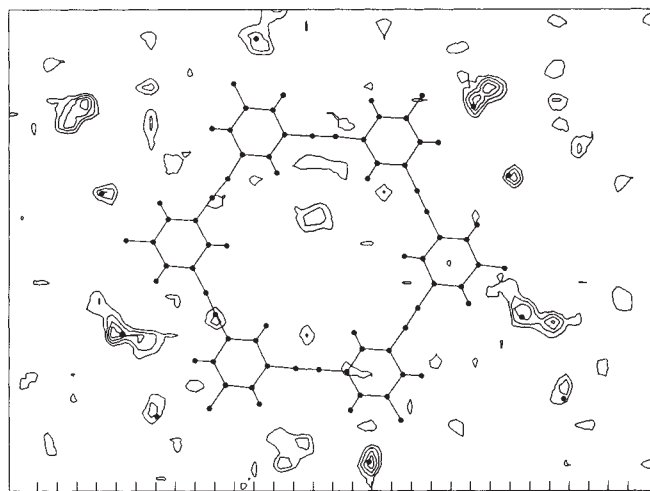


FIG. 2 Difference Fourier diagram for **1** (projected along the mean plane passing through atoms of the macrocycle) before the solvent molecules were added to the model. The peaks in the difference Fourier map coincide with the positions of the oxygen atoms of the solvents in the final model.

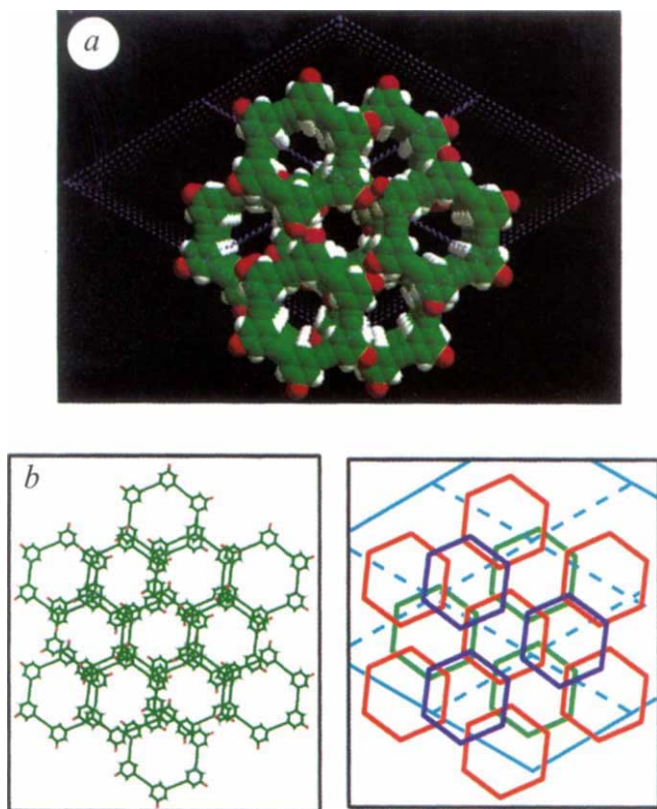


FIG. 3 a, Space-filling rendition of **1** viewed along the [010] axis. Extended channels with a diameter of  $\sim 9$  Å are clearly visible. b, Packing diagram of three layers viewed along the [010] axis (left) and a schematic showing their relative arrangement (right). With reference to Fig. 1, the central hexagon in the middle layer (red) has a total of 12 nearest neighbours (6 in the same layer, 3 in the upper layer (blue), and 3 in the lower layer (green). The layers are stacked in the sequence ... ABCABC ... of cubic closest packing.

Waals distances (Fig. 3b). If we were to represent each molecule as a sphere located at the centre of the macrocycle, the structure reduces to that of a face-centred cubic arrangement. Because of the large holes within the macrocycles, the framework atoms only occupy about half the total cell volume. The remaining void space is largely filled with solvent molecules.

The synthesis of new macrocycles opens up a number of possibilities. Because sections of the desired channel are pre-assembled with covalent bonds before crystallization, the size and shape of the channel's cross-section are dictated by the macrocycle's inner geometry. Also, as the organic building blocks can be readily modified, one can in principle tailor the microscopic cavity environment. Work is in progress to incorporate metal-coordinating functional groups to assemble binary organic-inorganic networks with concomitantly stronger intermolecular bonds. □

Received 16 May; accepted 15 September 1994.

- Ozin, G. A., Kuperman, A. & Stein, A. *Angew. Chem. int. Edn engl.* **28**, 359–376 (1989).
- Rolinson, D. R. *Chem. Rev.* **90**, 867–878 (1990).
- Zerkowski, J. A., MacDonald, J. C., Seto, C. T., Wierda, D. A. & Whitesides, G. M. *J. Am. chem. Soc.* **116**, 2382–2391 (1994).
- Chang, Y. L., West, M. A., Fowler, F. W. & Lauher, J. W. *J. Am. chem. Soc.* **115**, 5991–6000 (1993).
- Lehn, J.-M., Mascal, M., DeCian, A. & Fischer, J. J. *chem. Soc., Perkin Trans. 2* 461–467 (1992).
- García-Tellado, F., Geib, S. J., Goswami, S. & Hamilton, A. D. *J. Am. chem. Soc.* **113**, 9265–9269 (1991).
- Etter, M. C. & Admond, D. A. *J. chem. Soc., chem. Commun.* 589–591 (1990).
- Michaelides, A., Kiritsis, V., Skoulitsa, S. & Aubury, A. *Angew. Chem. int. Edn engl.* **32**, 1495–1497 (1993).
- Simard, M., Su, D. & Wuest, J. D. *J. Am. chem. Soc.* **113**, 4696–4698 (1991).
- Ermer, O. & Eling, A. *Angew. Chem. int. Edn engl.* **27**, 829–833 (1988).
- Ermer, O. & Lindenberg, L. *Helv. chim. Acta* **74**, 825–877 (1991).
- Hoskins, B. F. & Robson, R. *J. Am. chem. Soc.* **112**, 1546–1554 (1990).
- Duchamp, D. J. & Marsh, R. E. *Acta Crystallogr.* **B25**, 5–19 (1969).
- Abrahams, B. F., Hoskins, B. F., Michail, D. M. & Robson, R. *Nature* **369**, 727–729 (1994).
- Davies, J. E., Kemula, W., Powell, H. M. & Smith, N. O. *J. incl. Phenom.* **1**, 3–44 (1983).
- Weber, E. *Molecular Inclusion and Molecular Recognition: Clathrates I* (ed. Weber, E.) 1 (Topics in Current Chemistry Vol. 140, Springer, Berlin, 1987).
- Abbott, S. J. et al. *J. chem. Soc., chem. Commun.* 796–797 (1982).
- Weber, E., Pollex, R. & Czujler, M. *J. org. Chem.* **57**, 4068–4070 (1992).
- Ghadiri, M. R., Granja, J. R., Milligan, R. A., Moore, D. E. & Khazanovich, N. *Nature* **366**, 324–327 (1993).
- Zhang, J., Pesak, D. J., Ludwick, J. L. & Moore, J. S. *J. Am. chem. Soc.* **116**, 4227–4239 (1994).
- Zhang, J. & Moore, J. S. *J. Am. chem. Soc.* **114**, 9701–9702 (1992).
- Zhang, J. & Moore, J. S. *J. Am. chem. Soc.* **116**, 2655–2656 (1994).
- Sheldrick, G. M. *Crystal Solution Program* (Inst. fuer Anorg. Chemie, Göttingen, 1993).
- Desiraju, G. R. *Crystal Engineering* 92–101 (Elsevier, New York, 1989).
- Etter, M. C. *Acc. chem. Res.* **23**, 120–126 (1990).
- Boeyens, J. C. A. & Pretorius, J. A. *Acta crystallogr.* **B33**, 2120–2124 (1977).
- Greenwood, N. N. & Earnshaw, A. *Chemistry of the Elements* 304 (Pergamon, Oxford, 1986).

SUPPLEMENTARY INFORMATION. Requests for a complete set of atomic coordinates and crystallographic details should be addressed to Mary Sheehan at the London editorial office of *Nature*.

ACKNOWLEDGEMENTS. We thank J. W. Kampf for collecting the data for this crystal and S. R. Wilson for a copy of the SHELXL-93 program. J.S.M. thanks the US NSF, the NSF Young Investigator programme (1992–97) and the 3M company (non-tenured faculty awards programme) for their support. S.L. thanks the A. P. Sloan Foundation, and the J. D. and C. T. MacArthur Foundation for fellowships.

to optimize hydrogen bonding. The stack mode further benefits from efficient van der Waals contacts and electrostatic interactions between aromatic rings in adjacent layers (see below). Each hydrogen-bonded motif (without the solvents) in a sheet could be described according to Etter's terminology as an  $R_3^3$  (32) graph set<sup>25</sup>. The O–O distance between two hydrogen-bonded phenols is 2.69 Å and the angle C–O–O is 117.2°. If we were to assume a geometrically calculated position for the hydrogen atoms, whereby the C–O–H angle is 120° (C–O–H angle in hydroquinone by neutron diffraction was found to be 121.0°), then the O–H...O angle of the hydrogen bond is 175.4°, very close to the optimum value of 180° (ref. 26). To achieve this optimum hydrogen bonding, there is a slight twist in the macrocycles with respect to the *a* axis. Each sheet has two types of holes: (1) the 9-Å hole generated by the macrocycle and (2) the 9.3-Å hole generated by the hydrogen bonds. Each of these cavities is filled with solvent molecules which hydrogen bond to the macrocycles and with one another. Methyl and methylene groups from the solvent molecules fill space in the cavities of adjacent layers, as well as in the layer from which they originate. The sheets are separated by 3.33 Å and stack in an ...ABCABC... sequence of cubic closest packing, identical to that of  $\beta$ -graphite<sup>37</sup>. This stacking aligns the macrocycles in such a way as to generate an extended channel 9 Å in diameter running parallel to the *c* axis (Fig. 3a). Proper registry between macrocycles in adjacent layers is crucial for extended channels. This alignment seems to be achieved by van der Waals and electrostatic interactions between aromatic rings of adjacent layers. The slight lateral offset between aromatic rings seen in Fig. 3b is a common feature of the stack pattern and has been attributed to optimization of electrostatic interactions (minimization of  $\pi$ – $\pi$  electron repulsion and maximization of the interaction of positively charged hydrogen atoms of one ring with the negatively charged  $\pi$ -electrons of an adjacent ring)<sup>24</sup>. Each macrocycle has six nearest neighbours in a layer and three from each of the adjacent layers, making a total of 12 macrocycles at van der



# Direct demonstration of Heisenberg's uncertainty principle in a superconductor

W. J. Ellison, M. Matters, U. Geigenmüller & J. E. Mooij

Department of Applied Physics, Delft University of Technology and Delft Institute for Micro-Electronics and Submicron Technology (DIMES), PO Box 5046, 2600 GA Delft, The Netherlands

A HEISENBERG uncertainty relation exists between any two non-commuting variables of a quantum-mechanical system. In a superconductor, two such variables are the number,  $n$ , of Cooper pairs and the phase,  $\phi$ , of the superconducting wavefunction. Suppressing fluctuations in either variable should lead to enhanced fluctuations in the other<sup>1,2</sup>. To demonstrate this effect, we have fabricated a structure in which the quantum-mechanical fluctuations in the phase of a superconducting grain can be suppressed. We measure the supercurrent that flows through two Josephson tunnel junctions of small capacitance that are connected to the grain. The capacitance of the grain is itself so small that the number of Cooper pairs is well defined—charge transport through the grain is possible only through quantum-mechanical fluctuations in  $n$ . The phase of the grain is coupled to a large superconducting reservoir such that the fluctuations in  $\phi$  can be controllably suppressed. The enhanced fluctuations in  $n$  that result from this coupling give rise to a large increase in the supercurrent through the grain.

In the superconducting state, Cooper pairs of electrons form a condensate that can be described with one macroscopic wavefunction. As a general consequence of this description, the number of Cooper pairs and the phase of the superconducting wavefunction are non-commuting variables<sup>1</sup>. The corresponding Heisenberg uncertainty relation has the form  $\Delta n \Delta \phi \geq 1/2$

(closely connected with the uncertainty relation between charge and magnetic flux  $\Delta Q \Delta \Phi \geq \hbar/2$ ). Because of the periodic nature of  $\phi$ , slight modifications are necessary when  $\Delta \phi$  becomes of the order of  $2\pi$  (ref. 3). Anderson<sup>2</sup> discussed the implications of this uncertainty relation in connection with the Josephson effect<sup>4</sup>. For an isolated grain the number of Cooper pairs is fixed and the phase is completely uncertain. If two macroscopic superconductors are connected via a thin insulating barrier, the phase difference between the two is well-determined. The number of Cooper pairs in each superconductor fluctuates strongly. Depending on the phase difference, charge will flow through the barrier. With present-day nanofabrication technology it is possible to study experimentally the intermediate regime where the uncertainties in  $n$  and  $\phi$  are comparable. An example is the system of two small Josephson junctions in series. The capacitance of the junctions can be made so small that the electrostatic energy required to add one Cooper pair to the island between the junctions is much larger than the thermal energy  $k_B T$ . This charging energy will tend to keep the number of Cooper pairs on the island fixed. The Josephson coupling connects charge eigenstates differing by one Cooper pair. This coupling induces quantum-mechanical fluctuations of the charge, which result in a supercurrent through the device. The influence of the ratio of the charging energy and the Josephson coupling energy on the maximum supercurrent of a double junction has been studied both theoretically<sup>5,6</sup> and experimentally<sup>7–11</sup>.

We have used a different approach: we keep the junction parameters of the double junction fixed and explicitly use the fact that, due to the Heisenberg relation, charge fluctuations increase when fluctuations in the phase are suppressed. Suppression of the phase fluctuations can be achieved by coupling the island through a third Josephson junction to a superconducting reservoir that has a large capacitance to ground. Because the reservoir is otherwise unconnected, the average current through this junction must be zero and the classical phase of the superconducting reservoir is equal to the average value of the phase  $\phi$  of the island. Quantum fluctuations of  $\phi$  will give rise to a

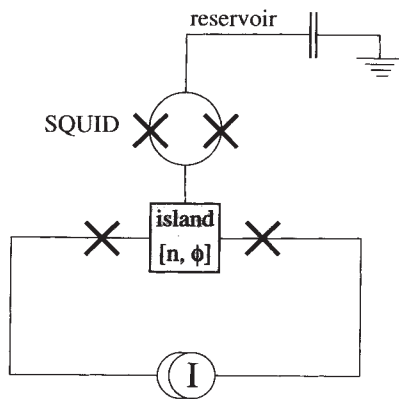


FIG. 1 Layout of our device. The crosses are Al-Al<sub>2</sub>O<sub>3</sub>-Al Josephson junctions that are connected with aluminium lines. The junction parameters are such that the quantum-mechanical fluctuations of the number of Cooper pairs and the phase of the island are comparable. The large capacitor represents the capacitance to ground of a macroscopic lead that is not connected to any measuring apparatus. This reservoir has a superconducting phase that is constant in time. The SQUID, consisting of two junctions in parallel, can be treated as a single junction with a flux-dependent Josephson coupling energy. The stronger the effective Josephson coupling between the island and the reservoir, the more the fluctuations in the phase of the island are suppressed. All leads are filtered with feedthrough filters at room temperature, and with copper powder filters at mixing-chamber temperature. The resulting frequency-dependent response of the bias circuitry is responsible for a non-ideal behaviour of the current source.

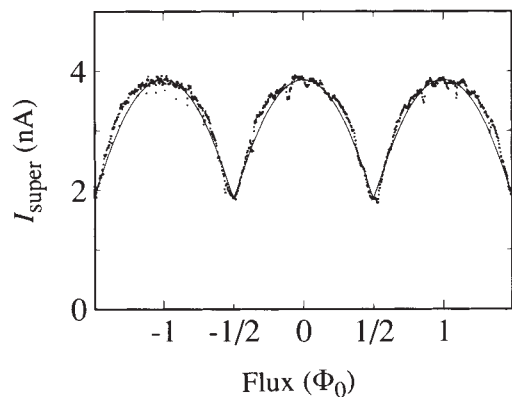


FIG. 2 Maximum supercurrent versus flux through the SQUID. The dots represent the switching current measured through the junctions connected to the current source. The solid line shows the calculated results, scaled linearly to account for the small but finite classical noise in the leads and the environmental impedance in the measurement. Flux is plotted in units of the flux quantum  $\Phi_0 = h/2e$ . For zero flux in the SQUID the coupling between the phase of the island  $\phi$  and the reservoir is strongest and fluctuations in  $\phi$  are suppressed. As a consequence, fluctuations in the number of Cooper pairs are enhanced which results in a large increase of the supercurrent. The magnetic field, applied perpendicular to the device, is of the order of a few gauss and does not influence the other junctions of the system. The measurements are performed at a temperature of 10 mK. We have verified that there is no non-integral offset charge present on the island at the time of the measurement.

time-dependent phase difference. As the Josephson energy of the junction is increased by such a phase difference, fluctuations in  $\phi$  will be suppressed. Stronger effective Josephson coupling results in stronger suppression of the fluctuations. To be able to actually tune fluctuations in the phase, we have connected the island to a superconducting reservoir via two junctions in parallel, a SQUID (Fig. 1). We use this SQUID as a single junction of which the critical current, and thereby the effective Josephson coupling energy, can be changed by applying a small magnetic field perpendicular to the device<sup>11,12</sup>. The effective Josephson coupling is largest when the number of magnetic flux quanta threading the loop is integral. For half-integral numbers, the Josephson coupling will be close to zero so that the island is practically decoupled from the superconducting reservoir.

We have fabricated the device shown in Fig. 1, choosing junction parameters such that the Josephson coupling energy is of the same order of magnitude as the charging energy. In this regime both the phase of the island and the number of Cooper pairs are quantum-mechanical variables that fluctuate around their equilibrium value. We have measured the current-voltage characteristics of the two junctions that are connected to the current source. We observe a supercurrent branch with a residual resistance that is lower than our measuring sensitivity of 10  $\Omega$ . We can clearly identify a switching current at which the voltage jumps to the superconducting gap. In Fig. 2 we plot the measured switching current as a function of flux through the SQUID. We observe a large modulation that is periodic with a period of one flux quantum through the SQUID. For integral number of flux quanta phase fluctuations are most strongly suppressed and the resulting enhanced charge fluctuations lead to an increase in the supercurrent.

We have calculated the influence of the flux on the supercurrent through the device using the method described in ref. 5. The ground-state energy is calculated from the electrostatic hamiltonian which is the sum of the charging energy of the island and the Josephson coupling energy of all four junctions. The critical current of the device follows from the derivative of the ground-state energy with respect to the phase difference of the measuring contacts. Thermal fluctuations of  $\phi$  at our measuring temperature of 10 mK are negligible, considering the energy difference of 2 K between the ground state and the first excited state. The influence of the flux through the SQUID on the critical current is the purely quantum-mechanical consequence of tuning the phase fluctuations on the island and vanishes in the classical limit where the Josephson coupling energy is much larger than the charging energy. The calculated critical current is generally found to be larger than the measured switching current because of small but finite classical noise present in the leads<sup>9,10,13</sup>. The exact relation depends on the details of the environmental impedance. In absence of a generally established theory we have scaled the critical current linearly. Figure 2 shows good agreement between our results and the theoretically predicted values of both relative amplitude and periodicity of the maximum supercurrent.

We have shown experimentally that fluctuations in the phase of the superconducting island of a double-junction system can be modulated without changing the Josephson coupling energy of the junctions that carry the bias current. The resulting modulation of the supercurrent clearly illustrates the Heisenberg uncertainty principle for phase and number of Cooper pairs on a superconductor. Squeezing the fluctuations of one variable by coupling to a reservoir, to enhance the transport connected with the other conjugate variable, may find wider application in quantum coherent circuits.  $\square$

4. Josephson, B. D. *Phys. Lett.* **1**, 251–253 (1962).
5. Averin, D. V. & Likharev, K. K. in *Mesoscopic Phenomena in Solids* (eds Althuler, B. L., Lee, P. A. & Webb, R. A.) 173–271 (Elsevier, Amsterdam, 1991).
6. Matveev, K. A., Gissel-fält, M., Glazman, L. I., Jonson, M. & Shekter, R. I. *Phys. Rev. Lett.* **70**, 2940–2943 (1993).
7. Geerligs, L. J., Anderregg, V. F., Romijn, J. & Mooij, J. E. *Phys. Rev. Lett.* **65**, 377–380 (1990).
8. Tuominen, M. T., Hergenrother, J. M., Tighe, T. S. & Tinkham, M. *Phys. Rev. Lett.* **69**, 1997–2000 (1992).
9. Joyez, P., Lafarge, P., Flipe, A., Esteve, D. & Devoret, M. H. *Phys. Rev. Lett.* **72**, 2458–2461 (1994).
10. Eiles, T. M. & Martinis, J. M. *Phys. Rev.* **B50**, 627–630 (1994).
11. Haviland, D. B., Harada, Y., Deising, P., Chen, C. D. & Claeson, T. *Phys. Rev. Lett.* **73**, 1541–1544 (1994).
12. Fulton, T. A., Gammel, P. L., Bishop, D. J. & Dunkleberger, L. N. *Phys. Rev. Lett.* **63**, 1307–1310 (1989).
13. Kautz, R. L. & Martinis, J. M. *Phys. Rev.* **B42**, 9903–9937 (1990).

ACKNOWLEDGEMENTS. We thank Yu. V. Nazarov for helpful discussions. This work was supported by the Netherlands Foundation for Fundamental Research on Matter (FOM).

## Effect of ozone depletion on atmospheric CH<sub>4</sub> and CO concentrations

S. Bekki, K. S. Law & J. A. Pyle

Centre for Atmospheric Science, Department of Chemistry, Lensfield Road, University of Cambridge, Cambridge CB2 1EW, UK

GLOBAL surface-based measurements of atmospheric methane and carbon monoxide concentrations revealed a marked and unexpected decrease in their growth rates in 1991 and 1992, particularly in the Northern Hemisphere<sup>1,2</sup>. Changes in emissions are unlikely to be the sole reason for the sudden reduction in the concentrations of these source gases<sup>2,3</sup>. The unprecedentedly large depletion of stratospheric ozone observed in 1991 and 1992 (ref. 4) may have contributed to the sharp decrease in the growth rates of both CH<sub>4</sub> and CO by exposing the troposphere to more ultraviolet radiation. This would have resulted in increased concentrations of the hydroxyl radical, OH<sup>•</sup>, which is the major atmospheric sink for both CH<sub>4</sub> and CO. Here we present two-dimensional model simulations which allow us to assess the significance of the link between stratospheric ozone depletion and the observed trends of CH<sub>4</sub> and CO. We find that the low values in stratospheric ozone concentration can account for almost half of the 1992 decrease in the CH<sub>4</sub> and CO growth rates.

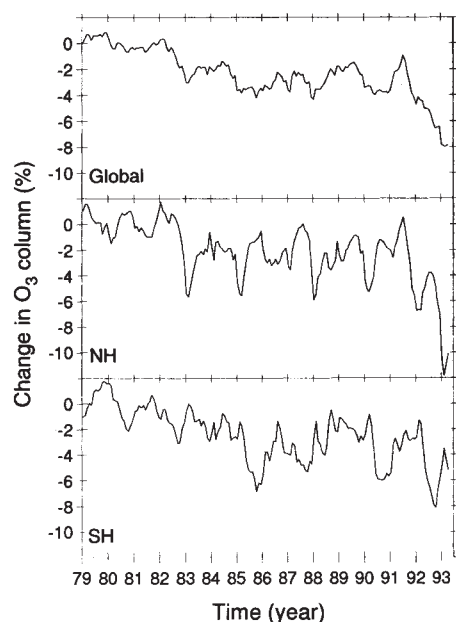
Satellite observations, combined with balloon-borne and ground-based measurements, have shown that global stratospheric ozone has significantly decreased during the last decade with the main reduction taking place in the lower stratosphere<sup>5</sup>. In 1991, global ozone started to decrease at a much faster rate, reaching a record low during the winter and spring of 1992–93 (ref. 4). Stratospheric ozone losses should have been accompanied by increases in ultraviolet radiation reaching the troposphere, all other factors being constant. Enhanced UV-B radiation in the troposphere accelerates the photodissociation of ozone to electronically excited atomic oxygen, O(<sup>1</sup>D), and therefore increases OH<sup>•</sup> concentrations<sup>6–9</sup>. The removal of most tropospheric trace gases, including CH<sub>4</sub> and CO, is dominated by reaction with OH<sup>•</sup>. Therefore, any changes in tropospheric OH<sup>•</sup> resulting from stratospheric ozone changes should be accompanied by changes in the growth rates of trace gases.

Surface measurements from global networks have revealed a decline in the rate of increase of CH<sub>4</sub> during the last decade, most noticeably at the end of the 1980s<sup>5,10,11</sup>. The growth rate decreased from 17–21 parts per billion by volume (p.p.b.v.) yr<sup>-1</sup> (~1.2% yr<sup>-1</sup>) at the beginning of 1980s to 12–14 p.p.b.v. yr<sup>-1</sup> (~0.7% yr<sup>-1</sup>) during 1988–90 (ref. 5). Unexpectedly, a sharp reduction in the CH<sub>4</sub> growth rate has been reported for 1992 (ref. 1). Most of this reduction occurred in the Northern Hemi-

Received 16 June; accepted 21 September 1994.

1. Tinkham, M. *Introduction To Superconductivity* (McGraw-Hill, New York, 1975).
2. Anderson, P. W. in *Progress in Low Temperature Physics* Vol. V (ed. Gorter, C. J.) 1–43 (North-Holland, Amsterdam, 1967).
3. Carruthers, P., Nieto, M. M. *Rev. mod. Phys.* **40**, 411–440 (1968).





sphere. For the first time in this 10-yr record, the  $\text{CH}_4$  growth rate in the Northern Hemisphere dropped from 12 p.p.b.v.  $\text{yr}^{-1}$  in summer 1991 to below zero by the end of 1992 (ref. 1).

$\text{CO}$  measurements indicated a small positive trend between 1981 and 1986 followed by a decrease between 1987 and 1992<sup>12</sup>. Most recently, global measurements made from 1990 to 1993 showed a more pronounced decrease in  $\text{CO}$  at an average rate of about 7 p.p.b.v.  $\text{yr}^{-1}$  ( $\sim 6\% \text{ yr}^{-1}$ ) in the Northern Hemisphere and 4 p.p.b.v.  $\text{yr}^{-1}$  ( $\sim 7\% \text{ yr}^{-1}$ ) in the Southern Hemisphere<sup>2</sup>.

We use a global two-dimensional model of the atmosphere together with the Total Ozone Mapping Spectrometer (TOMS) monthly mean dataset to assess the contribution of stratospheric ozone changes to variations in the growth rates of  $\text{CH}_4$  and  $\text{CO}$ . In order to predict changes in the penetration of ultraviolet radiation consistent with observed stratospheric ozone changes, photolysis rates in the troposphere are calculated by scaling the model ozone columns with the corresponding TOMS ozone columns. The model is run for two cases from January 1965 to April 1993. The control run uses a perpetual annual ozone column cycle calculated by averaging the TOMS data over the 3-yr period November 1978 to October 1981. The perturbation run uses the perpetual annual cycle before November 1978 followed by actual TOMS data up to April 1993.

Global ozone decreased by  $\sim 4\%$  between 1979 and 1985 (Fig. 1), followed by a small recovery during the years 1986–90 (ref. 13) and finally a further drop after the volcanic eruption of Mount Pinatubo in June 1991. The average ozone column decreased by  $\sim 10\%$  in the Northern Hemisphere from mid-1991 to the beginning of 1993. Modelled changes in surface ultraviolet flux at 300 nm are in good agreement with trends deduced from spectral measurements made at Toronto ( $44^\circ \text{N}$ ) from 1989 to 1993 (ref. 14). This suggests that scattering by volcanic aerosols played a minor role in the variations of surface UV-B flux following the eruption of Mount Pinatubo in 1991 (ref. 15). The modelled trend in global  $\text{OH}^\bullet$  due to the ozone column changes (Fig. 2) is approximately  $+0.2\text{--}0.3\% \text{ yr}^{-1}$  between 1979 and 1990. However, from mid-1991 to spring 1993, global  $\text{OH}^\bullet$  increases by 6%. The increase is smaller in the Southern Hemisphere (5%) than in the Northern Hemisphere (8%) and peaks in autumn 1992, much earlier than in the Northern Hemisphere. The modelled change in  $\text{OH}^\bullet$  is more marked in 1992 than during the 1980s, and so should have been less likely to be masked by variations in other factors governing  $\text{OH}^\bullet$  such as emissions of  $\text{CH}_4$ ,  $\text{CO}$  or nitrogen oxides.

FIG. 1 Time series of the percentage difference in global and hemispheric average ozone columns between the control and perturbation runs (NH, Northern Hemisphere; SH, Southern Hemisphere). The model used in this study contains a detailed treatment of the  $\text{O}_x$ ,  $\text{NO}_x$ ,  $\text{HO}_x$ ,  $\text{ClO}_x$ ,  $\text{BrO}_x$ ,  $\text{CHO}_x$  and non-methane hydrocarbon (NMHC) chemistry<sup>19,20</sup> with recent kinetic data<sup>21</sup>. A new photolysis scheme, particularly appropriate for tropospheric studies has been included in the model<sup>22</sup>. The boundary conditions for  $\text{CO}_2$ , nitrous oxide ( $\text{N}_2\text{O}$ ) and CFCs are from a recommended scenario<sup>23</sup>. Emissions of  $\text{CH}_4$ ,  $\text{CO}$ ,  $\text{NO}_x$  and NMHCs are kept constant.

The differences between the runs in the growth rates of  $\text{CH}_4$  and  $\text{CO}$  are plotted against time in Figs 3 and 4 respectively. Any variation in the difference of the growth rates between the two runs can be treated as a change in the growth rate itself resulting from variations in tropospheric ultraviolet flux. The variation of the global  $\text{CH}_4$  growth rate follows the long-term variation of global ozone (Fig. 1): an overall decrease from zero growth in 1979 to  $-3 \text{ p.p.b.v. yr}^{-1}$  in 1985, a recovery from 1986 to 1989 and a decrease from a slightly positive growth rate in mid-1991 to  $-5 \text{ p.p.b.v. yr}^{-1}$  in autumn 1992. These results show that stratospheric ozone changes may have contributed to the slow decline in the  $\text{CH}_4$  growth rate before 1985. But in the late 1980s, a slight recovery in global ozone may have prevented an even more pronounced decrease in the  $\text{CH}_4$  growth rate. The most striking feature in the  $\text{CH}_4$  growth rate is a marked decrease of 7 p.p.b.v.  $\text{yr}^{-1}$  in the Northern Hemisphere from spring 1991 to autumn 1992. This indicates that the record-low ozone col-

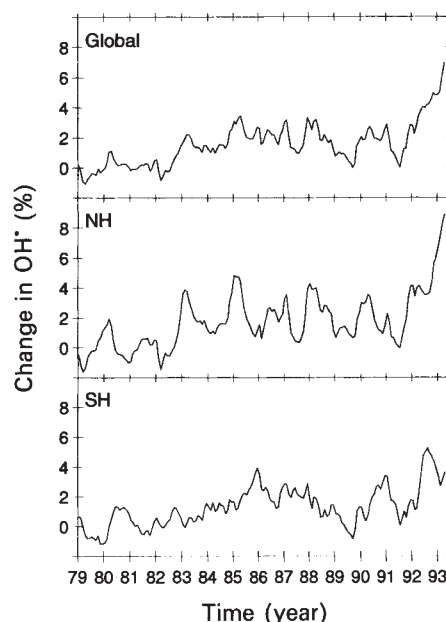


FIG. 2 Same as Fig. 1, but for  $\text{OH}^\bullet$  at the surface level ( $\sim 800 \text{ mb}$ ).

umn values in 1992 may have accounted for almost half of the large reduction in  $\text{CH}_4$  growth rate observed in 1992 (ref. 1). Reductions in  $\text{CH}_4$  emissions from gas fields and possibly biomass burning, which are not included in our simulations, may also have contributed to the decrease in the  $\text{CH}_4$  growth rate<sup>1,3</sup> in the Northern Hemisphere.

The variations in the CO growth rate (Fig. 4) show well defined interannual variations of  $\sim 2$  p.p.b.v.  $\text{yr}^{-1}$  which oscillate around zero, particularly in the Northern Hemisphere. As for  $\text{CH}_4$ , the global CO growth rate decreases sharply from 1991 to 1992. In the Northern Hemisphere, the drop in the CO growth rate translates into a decrease of 5–6 p.p.b.v. in CO from spring 1991 to spring 1993. This represents up to 40% of the abrupt decrease observed in the Northern Hemisphere during same period. But the observations show that the CO decrease began earlier, in 1990<sup>2</sup> or even 1987<sup>12</sup>. It is possible that reduced levels of anthropogenic emissions of  $\text{CO}^{16}$  (and non-methane hydrocarbons, NMHCs, which are oxidized to CO) and decreases in tropical biomass burning<sup>12</sup> may have contributed to the declining CO trend.

The trends of several other source gases which are removed by  $\text{OH}^\bullet$  are also found to be affected. For example, the growth rates of  $\text{CH}_3\text{CCl}_3$  and  $\text{CH}_3\text{Br}$  decrease by 1 part per trillion ( $10^{12}$ ) by volume (p.p.t.v.)  $\text{yr}^{-1}$  ( $\sim 0.6\%$   $\text{yr}^{-1}$ ) and by 1 p.p.t.v.  $\text{yr}^{-1}$  ( $\sim 2\%$   $\text{yr}^{-1}$ ), respectively from spring 1991 to autumn 1992 as a result of stratospheric ozone depletion.

The evidence presented here, in conjunction with the recent  $\text{CH}_4$  and CO observations<sup>1,2</sup>, suggest strongly that stratospheric ozone can have a significant effect on the removal of trace gases. This may be of importance in the future. If, as expected from international protocols on chlorofluorocarbon (CFC) emissions, stratospheric ozone levels begin to return to pre-CFC levels this will lead to faster rates of increase of source gases including  $\text{CH}_4$  and hydrofluorocarbons (HFCs). Our results also have implications for assessing the total radiative forcing from stratospheric ozone depletion. For example, the direct radiative forcing from a sustained stratospheric ozone depletion of 3% in

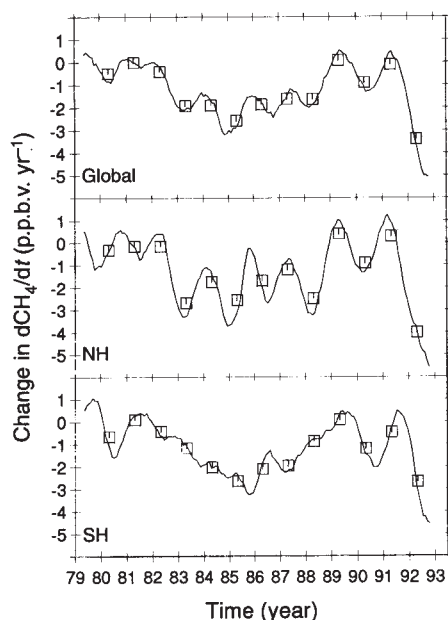


FIG. 3 Time series of the difference between the runs in the growth rates of global and hemispheric average  $\text{CH}_4$  mixing ratios at the surface level ( $\sim 800$  mb). Growth rates are calculated from the deseasonalized monthly means using the method of 'linear moving slopes' described in ref. 11 with a time span of 1 yr. Squares are 1-yr averages centred on April.

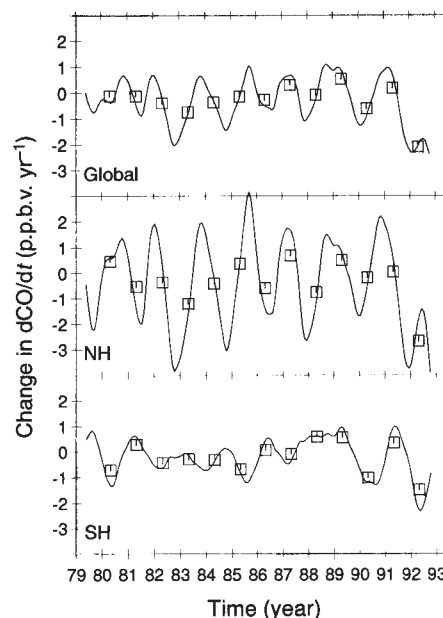


FIG. 4 Same as Fig. 3 except for CO.

global ozone has been estimated to be negative ( $-0.08$  to  $-0.09 \text{ W m}^{-2}$ ; ref. 17). In view of our results, this should also be accompanied by a reduction in  $\text{CH}_4$  of 3% and in tropospheric ozone of  $\sim 1\%$  (ref. 9) all other factors being constant. These perturbations represent an indirect forcing of  $-0.03$  to  $-0.04 \text{ W m}^{-2}$  with  $-0.02 \text{ W m}^{-2}$  from the  $\text{CH}_4$  change and  $-0.01$  to  $0.02 \text{ W m}^{-2}$  from tropospheric ozone change. This results in a total (direct+indirect) radiative forcing 30–50% higher than the direct forcing. Note that this climate feedback will have been obscured by increasing anthropogenic emissions during the past decade.

Interestingly, recent measurements indicate that ozone concentrations started to recover in the Northern Hemisphere in late 1993 (ref. 18). This may have translated into a return to higher growth rates for  $\text{CH}_4$  and CO. Measurements in 1994 will provide an additional test of the link between stratospheric ozone and the oxidizing capacity of the troposphere.  $\square$

Received 11 April; accepted 12 August 1994.

1. Dlugokencky, E. J. et al. *Geophys. Res. Lett.* **21**, 45–48 (1994).
2. Novelli, P. C. et al. *Science* **263**, 1587–1590 (1994).
3. Rudolph, J. *Nature* **368**, 19–20 (1994).
4. Gleason, J. F. et al. *Science* **260**, 523–526 (1993).
5. *Scientific Assessment of Ozone Depletion: 1991* (Report No. 25, World Meteorological Organization Global Ozone Research and Monitoring Project, Geneva, 1992).
6. Liu, S. C. & Trainer, M. J. *Atmos. Chem.* **6**, 221–233 (1988).
7. Thompson, A. M., Stewart, R. W., Owens, M. A. & Herwehe, J. A. *Atmos. Environ.* **23**, 519–532 (1989).
8. Madronich, S. & Granier, C. *Geophys. Res. Lett.* **19**, 465–467 (1992).
9. Fuglestad, J. S., Jonson, J. E. & Isaksen, I. S. A. *Tellus* (in the press).
10. Steele, L. P. et al. *Nature* **355**, 313–316 (1992).
11. Khalil, M. A. K. & Rasmussen, R. A. *Chemosphere* **26**, 803–814 (1993).
12. Khalil, M. A. K. & Rasmussen, R. A. *Nature* **370**, 639–641 (1994).
13. Herman, J. R., McPeters, R., Stolarski, R., Larko, D. & Hudson, R. J. *Geophys. Res.* **96**, 17297–17305 (1991).
14. Kerr, J. F. & McElroy, C. T. *Science* **262**, 1032–1034 (1993).
15. Vogelmann, A. M., Ackerman, T. P. & Turco, R. P. *Nature* **359**, 47–49 (1992).
16. Houghton, J. et al. (eds) *Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 1990).
17. Ramaswamy, V., Schwarzkopf, M. D. & Shine, K. P. *Nature* **355**, 810–812 (1992).
18. Planet, W. G. et al. *Geophys. Res. Lett.* **21**, 205–208 (1994).
19. Law, K. & Pyle, J. A. *J. geophys. Res.* **98**, 18377–18400 (1993).
20. Law, K. & Pyle, J. A. *J. geophys. Res.* **98**, 18401–18412 (1993).
21. DeMore, W. B. et al. *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling* (Evaluation No. 10, NASA/JPL Publ. 92-1, Pasadena, California, 1992).
22. Hough, A. M. *J. geophys. Res.* **96**, 7325–7362 (1991).

ACKNOWLEDGEMENTS. This work was supported by the UK Department of the Environment and the UK Universities Global Atmospheric Modelling programme (UGAMP).



# Improved growth and survival of offspring of peacocks with more elaborate trains

Marion Petrie

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

THERE is considerable controversy about what females gain from mate choice in a lekking species in which males provide no obvious resources. Females may gain direct benefits such as safe copulations or increased fertility by mating with particular males, or they may gain indirect benefits for their offspring<sup>1-3</sup>. It is difficult to look for paternal effects on offspring performance because it is hard to control for any differences in the material and genetic contribution provided by the female and in the environment of the offspring during development; previous experiments have controlled for rearing environment but not maternal effects<sup>4</sup>. Here I report results from a controlled breeding experiment, with females allocated to males at random and all offspring reared under the same conditions, which show that the offspring of successful lek peacocks (*Pavo cristatus*) with the most elaborate trains grow and survive better under nearly natural conditions.

Observation at Whipsnade Park, UK, revealed that peahens show mate preferences in that they always reject some potential males as mates and mate with the male that has the most elaborate train<sup>5</sup>. The degree of train elaboration, measured as the total number of eye-spots visible in photographs, is a predictor of a male's mating success<sup>5,6</sup>. Moreover, experimental reduction of the number of eye-spots in the train between years results in a decline in mating success for individuals<sup>6</sup>. This is evidence that females base their preference on some aspect of train morphology, although these data do not distinguish which of several possible attributes of the train females prefer, such as the overall symmetry of eye-spots in the train or the total area of eye-spots. To understand why a female preference for elaborate trains is maintained in this population, I looked to see whether there were any paternal effects on the growth of offspring. I removed males of known status from Whipsnade, and kept them in separate pens before assigning four breeding females randomly to each pen. The resulting offspring were reared under the same conditions and their growth monitored at regular intervals (Fig. 1).

The design of the experiment was to ensure that the main difference between the pens was the attractiveness of the males, but it is possible that during the random assignment of females chance differences were created between pens in the mean size of the females and this may subsequently contribute to any differences found between pens in offspring size. However, there were no significant differences between pens in any of the measured variables taken on females (ANOVA analysis of variance: body weight  $F(7, 24)=0.32$ ,  $P=0.94$ ; tarsus length  $F(7, 24)=0.59$ ,  $P=0.7$ ; condition (weight/tarsus; ref. 3)  $F(7, 24)=0.57$ ,  $P=0.77$ ; spur length  $F(7, 24)=0.98$ ,  $P=0.47$ . Spur length of females increases with female age (M.P., unpublished results).

An analysis of variance was carried out on the size of the offspring at day 84 to see whether there was a significant effect of paternity on offspring growth. The potential confounding influences of hatch date and egg weight were removed by entering them as covariates in the analysis. There were significant differences between fathers in offspring weight, controlling for any differences due to hatch date or egg weight (main effect, father,  $F(7, 324)=6.57$ ,  $P<0.0001$ ; covariates: (1) hatch date,  $F(1, 324)=124.73$ ,  $P<0.0001$ ; (2) egg weight,  $F(1, 324)=0.051$ ,  $P>0.8$ ). Treating the offspring sexes separately, there were sig-

nificant differences between fathers in the growth of sons and daughters (daughters: main effect, father,  $F(7, 165)=3.59$ ,  $P<0.001$ ; covariates: (1) hatch date,  $F(1, 165)=62.11$ ,  $P<0.001$ ; (3) egg weight,  $F(1, 165)=1.12$ ,  $P>0.2$ . Sons: main effect, father,  $F(7, 146)=2.71$ ,  $P<0.01$ ; covariates: (1) hatch date,  $F(1, 146)=69.39$ ,  $P<0.001$ ; (2) egg weight,  $F(1, 146)=0.1$ ,  $P>0.7$ ).

A multiple regression analysis was carried out to see whether growth differences between offspring were dependent upon any of the measured morphological variables of the father. A preliminary analysis of the data revealed that three variables were positively and independently correlated with the size of offspring at day 84. These were, father's tarsus length, father's train length and the size of the eye-spots in the father's train. These variables were entered into the multiple regression together with the hatch date which has a strong effect on offspring growth (later-hatched chicks being smaller than earlier-hatched chicks) and egg weight

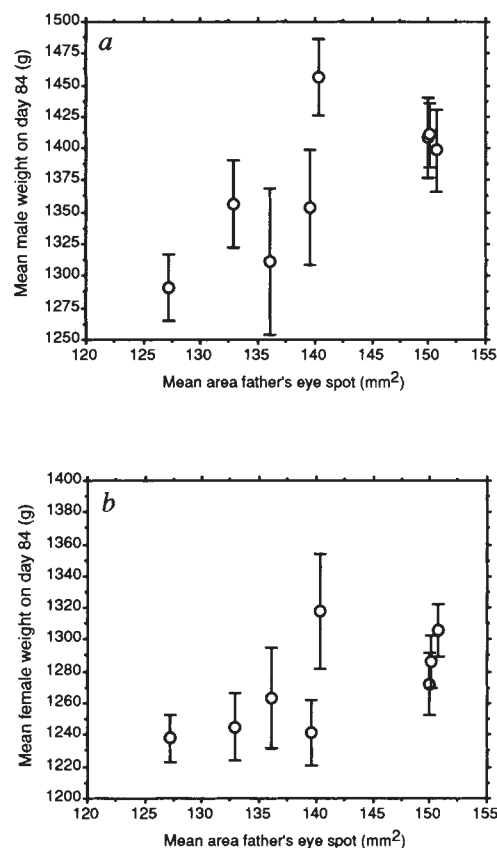


FIG. 1 The relationship between the mean weight ( $\pm$  standard errors) of offspring at day 84 and the size of the eye-spots in their father's train. a, This relationship for sons ( $n=8$ ,  $r=0.74$ ,  $P=0.035$ ), and b, daughters ( $n=8$ ,  $r=0.66$ ,  $P=0.077$ ) (all variables were log-transformed for calculation of regression statistics). Eight free-ranging full-trained displaying lek males whose mating success varied were removed from Whipsnade Park during February 1991 and transferred to a farm in Norfolk, UK. The peacocks were housed in separate pens and 4 naive adult peahens, known to be at least 2 years old, were measured and randomly assigned to each pen on 14 March. Pens were checked every day for eggs (eggs could not be assigned to a hen unless egg-laying was witnessed), which were labelled and removed. Groups of eggs originating from several different pens were placed under broody domestic chickens for incubation. Eggs were removed from the hens after 26 days and placed in a hatcher, where each egg had its own compartment. Hatched chicks were given an individual colour ring combination and were reared together under standard conditions with heat lamps and food and water *ad libitum*. Females produced 519 eggs and the growth of the surviving 349 offspring was monitored. Chicks were weighed and measured on days 0, 7, 14, 28, 42, 56 and 84.

TABLE 1 Results from a multiple regression analysis with the size of offspring at day 84 as the dependent variable and five independent variables ( $F(5, 319) = 28.49, P < 0.0001$ )

| Independent variable                        | $\beta$ | $T$    | Significance |
|---------------------------------------------|---------|--------|--------------|
| Mean area of 20 eye-spots in father's train | 0.16    | 3.41   | 0.0007       |
| Egg weight                                  | -0.025  | -0.54  | 0.59         |
| Hatch date                                  | -0.52   | -11.05 | 0.0001       |
| Father's tarsus length                      | 0.047   | 0.89   | 0.37         |
| Length of father's train                    | 0.094   | 1.74   | 0.084        |

Train measures were made on shed feathers. The fathers in the pens started to shed their upper tail coverts between July and August; all the shed feathers were collected for each male. Several measures were taken on the trains of each bird: the total number of feathers, the weight of the train, the length of the longest feather and the mean size of 20 eye-spots. The area of the blue centre of each eye-spot was measured using a digital pad. This multiple regression only includes those measures that were positively and independently correlated with offspring growth.  $\beta$ , Regression coefficient.

(a potentially confounding variable); results are shown in Table 1. Males with larger eye-spots in their trains tended to have larger offspring after controlling for the other four variables. Figure 1a and b shows the positive relationship between the mean size of offspring produced in each pen and the size of eye-spots in the father's train for sons and daughters respectively.

Although there were no differences in the mother's phenotype between pens, it is possible that females produced larger or better-quality eggs in response to being mated to males with more elaborate trains and that this could account for the differences in growth. But the weight of an egg does not predict the size of offspring at day 84 (Table 1) and is controlled for statistically in all analyses which show paternal effects on offspring growth. Although it is theoretically possible that some differential maternal provisioning of the eggs, independent of egg size,

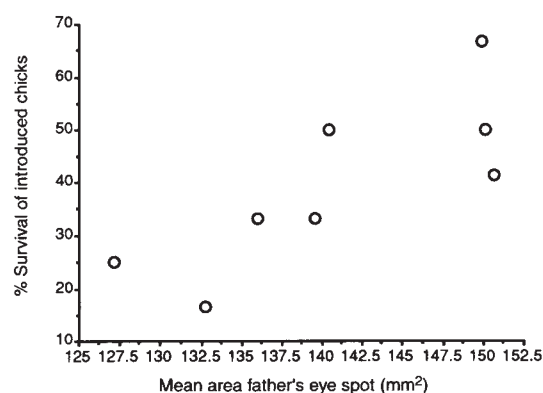


FIG. 2 The relationship between the proportion of offspring surviving introduction into Whipsnade Park and the size of eye-spots in their father's train ( $n=8, r=0.79, P=0.018$ ). In January and February 1992, 12 offspring (7 males and 5 females) from each of the 8 males (3 from each of the 4 females per male) were introduced into Whipsnade Park. A matched sample of young were chosen from each pen so that there was no overall significant differences in hatching dates or weights (at day 84,  $F(7, 95)=0.838, P=0.559$ ; at introduction  $F(7, 95)=0.358, P=0.924$ ) of the offspring between fathers. Care was taken to release the offspring in batches of eight, consisting of one young of the same sex from each pen. There was a significant relationship between the proportion of sons surviving the introduction and the size of eye-spots in the father's train ( $n=8, r=0.84, P=0.0094$ ) but not for the proportion of daughters surviving ( $n=8, r=0.596, P=0.119$ ), but the sample of daughters for each male was small and there is a trend in the same direction.

could be responsible for larger trained males producing larger offspring, there is no evidence in any species<sup>7</sup> that this sort of differential investment occurs. I tested to see whether there were any measurable differences in the quality of eggs laid for the different fathers in the following breeding season by measuring the lipid and protein content of two eggs laid by two different females for each male and found no significant differences between fathers in any of the measured variables (per cent lipid in yolk ( $n=16$ ), mean was  $64.73 \pm 1.22$  s.d., ANOVA  $F(7, 15) = 0.85, P=0.57$ ; per cent protein in albumen, mean was  $82.18 \pm 1.842$  s.d., ANOVA  $F(7, 15)=0.33, P=0.92$ ; per cent protein in yolk, mean was  $29.41 \pm 0.62$  s.d., ANOVA  $F(7, 15) = 1.28, P=0.36$ ).

I released a sample of the offspring reared at the farm into Whipsnade Park (see Fig. 2 for details). Of the offspring introduced, 41% were present after 24 months. Several cases of mortality were reported and the condition of the corpses suggested that these were caused by fox (*Vulpes vulpes*) predation; other missing birds were presumed dead. Birds that were larger at the time of introduction were more likely to have survived their first year at Whipsnade. The mean weight of males that were presumed dead was 3.34 kg (s.d. = 0.27), whereas that of surviving males was 3.54 kg (s.d. = 0.21),  $t = -3.12, P < 0.003$ ; for females, the mean weight of those presumed dead was 2.67 kg (s.d. = 0.24) whereas the mean weight of survivors was 2.85 kg (s.d. = 0.18),  $t = -2.27, P < 0.03$ . there was a significant positive correlation between the weight of the birds at introduction and their weight at day 84 ( $n=95, r=0.61, P < 0.0001$ ). The proportion of birds surviving for each of the eight fathers was dependent upon their mean weight at day 84 ( $n=8, r=0.73, P=0.042$ ) and upon the size of the eye-spots in their father's trains (Fig. 2); however, multiple regression showed that neither of these factors significantly affected survivorship independently of the other (eye-spot size, partial  $F=3.975, P=0.103$ ; mean weight, partial  $F=1.96, P=0.22$ ).

The results show the offspring of highly ornamented males tend to grow better and that these advantages translate into differences in the chance of their subsequent survival under almost natural conditions. These data provide support for the idea that females may be gaining good 'viability' genes for their offspring when mating with attractive males and extend the previous evidence from lekking species which shows that preferred males are more likely to survive<sup>8,9</sup>.

Norris<sup>4</sup> has shown that male great tits (*Parus major*) with larger breast stripes, a trait preferred by females, produce more surviving sons. Cross-fostering experiments were used to control for differences in rearing conditions, but not for the possibility that higher quality females mate with preferred males; assortative mating of this sort has been shown to occur<sup>10</sup>. In this study random allocation of several females to each male controls for maternal effects on offspring performance. Although the evidence suggests that females may gain indirect benefits in the form of viability advantages, it does not exclude the possibility that females may also be gaining other advantages from their choice. □

Received 23 May; accepted 1 September 1994.

- Kirkpatrick, M. T. & Ryan, M. J. *Nature* **350**, 33–38 (1991).
- Pomiankowski, A., Iwasa Y. & Nee, S. *Evolution* **45**, 1422–1430 (1991).
- Iwasa, Y., Pomiankowski, A. & Nee, S. *Evolution* **45**, 1431–1442 (1991).
- Norris, K. *Nature* **362**, 537–539 (1993).
- Petrie, M., Halliday, T. & Sanders, C. *Anim. Behav.* **41**, 323–331 (1991).
- Petrie, M. & Halliday, T. R. *Behav. Ecol. Sociobiol.* (in the press).
- Williams, T. D. *Biol. Rev.* **68**, 35–59 (1994).
- Alatalo, R. V., Hoglund, J. & Lundberg, A. *Nature* **352**, 155–156 (1991).
- Petrie, M. *Anim. Behav.* **44**, 585–586 (1992).
- Møller, A. P. *Proc. R. Soc. Lond. B* **245**, 179–182 (1991).

ACKNOWLEDGEMENTS. I thank R. Kock, P. Olney and N. Lindsay for their support and for permission to work at Whipsnade Park; Q. Spatt for allowing me to work at his farm; N. Parillon, A. Williams and D. Parrott for assistance in the field; A. P. Møller for help in measuring eye-spots; J. Newman and W. Blanckenhorn for statistical advice; A. P. Møller and P. Cotgreave for comments on the manuscript; and the Natural Environment Research Council for financial support.



# Mapping motor representations with positron emission tomography

J. Decety\*, D. Perani†, M. Jeannerod\*, V. Bettinardi‡, B. Tadary\*, R. Woods‡, J. C. Mazziotta‡ & F. Fazio†

\* Vision et Motricité, INSERM U94, 16 avenue du Doyen Lépine, F-69500 Bron, France

† INB-CNR, University of Milan, Scientific Institute H San Raffaele, Via Olgettina 60, 20132 Milano, Italy

‡ Brain Mapping Division, Department of Neurology, UCLA School of Medicine, Los Angeles, California 90024-1763, USA

**BRAIN activity was mapped in normal subjects during passive observation of the movements of an 'alien' hand and while imagining grasping objects with their own hand. None of the tasks required actual movement. Shifting from one mental task to the other greatly changed the pattern of brain activation. During observation of hand movements, activation was mainly found in visual cortical areas, but also in subcortical areas involved in motor behaviour, such as the basal ganglia and the cerebellum. During motor imagery, cortical and subcortical areas related to motor preparation and programming were strongly activated. These data support the notion that motor learning during observation of movements and mental practice involves rehearsal of neural pathways related to cognitive stages of motor control<sup>1-3</sup>.**

Cerebral activation was monitored in normal humans (six right-handed male subjects, 22–41 years old) using a positron emission tomograph. The subjects had given their informed consent. Changes in regional cerebral blood flow (rCBF) were measured using the intravenous radioactively labelled water ( $H_2^{15}O$ ) bolus technique.

Two activation conditions (movement observation and motor imagery) and one control condition (visual inspection) were used. In all conditions three-dimensional graspable objects (cylinders and spheres of different size (1–4 cm), colour and orientation) were presented at reaching distance, through a three-

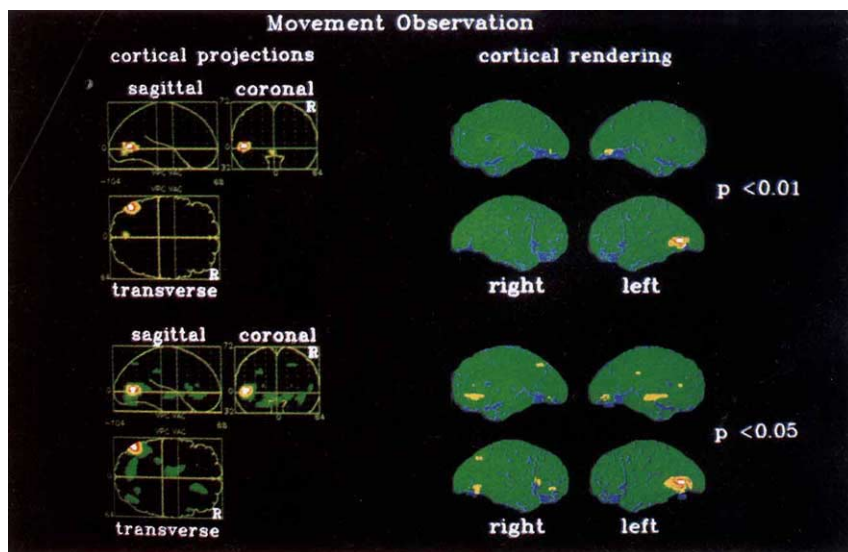
dimensional virtual reality system. Objects appeared in a random sequence at a rate of 0.3 Hz and at the same position in virtual space. During the movement observation condition, subjects were instructed to watch the movements of a virtual right hand grasping the objects 'as if it were their own hand'. There was a short training session during which the subjects wore a data glove on their right hand. This device allows real-time interaction between the image of the reconstructed hand and the objects in the virtual environment. During the motor imagery condition, subjects were instructed to 'imagine themselves grasping the objects with their right hand'. In this condition, no virtual hand was present. In both movement observation and motor imagery, no actual movement was produced as checked by electromyogram recordings. Finally, in the visual inspection condition, subjects had to inspect the objects without further instruction. This condition was selected to subtract from the others the baseline activity pertaining to the analysis of visual input and object recognition.

During the movement observation condition, the anterior cingulate cortex (areas 24 and 32) was found to be activated bilaterally. The highest cortical activation, however, was in the extrastriate visual areas, the medial occipital gyrus (A19) approaching inferior temporal area 37 and inferior parietal area 39. Although this rCBF increase was bilateral, it was more marked on the left. A unilateral left activation of the lingual gyrus (area 18) was also observed. Most of these areas are specialized for processing visual motion information<sup>4</sup>. Their activation may be explained by the processing of the moving virtual hand (this component was not present in the control visual inspection condition and was not subtracted). In addition, areas located more rostrally (such as area 37) might be homologous to those described in monkey inferotemporal cortex, where neurons fire during observation by the animal of hand-object interactions<sup>5</sup>. These neurons might encode goal-centred descriptions of motor actions, a function which is likely to be essential for observational learning.

At the subcortical level, movement observation was associated with activation of structures involved in motor behaviour. The lenticular nucleus and the substantia nigra in the basal ganglia complex were activated in the left hemisphere. There was also a bilateral (predominantly left) cerebellar activation (see Table 1 and Fig. 1).

**FIG. 1** Adjusted mean rCBF during observation of a moving hand grasping objects (MO), after subtraction of the visual inspection (control) condition. The data are displayed as statistical maps in three projections (sagittal, coronal and transverse) on the left, and on the right as cortical rendering onto the medial and lateral cortical surfaces of each hemisphere. The pixel values significantly ( $P < 0.01$  and  $P < 0.05$ ) higher during MO are shown in colour. The intercommissural line is set at zero on the sagittal and coronal projections. The vertical projections through the anterior commissure (VAC) and posterior commissure (VPC) are depicted on the sagittal and transverse projections. Distances are in millimetres above (+) and below (–) the intercommissural (ACPC) line, anterior (+) and posterior (–) to the VAC line and to the left (–) and the right (+) of the midline.

**METHODS.** The virtual reality system consists of a Virtual Research helmet connected to a Provision virtual reality transputer-based parallel calculator (Division Ltd, Bristol, UK) that combines a 3-D stereo image generator and 3-D tracking and gestural devices. All conditions were preceded by a short training session before the positron emission tomographic measurements. Instructions were given verbally 1 min before the scanning procedure. Presentation of the stimuli began 30 s before the bolus



injection. Absence of changes in muscular activity was verified throughout the experiment using surface electrodes placed on the biceps and triceps brachii and on finger extensors of the right upper limb.

TABLE 1 Coordinates and *P* values for regions in which there was more activation during movement observation

|                               | X   | Y   | Z   | MO-Control<br><i>P</i> | MO-MI<br><i>P</i> |                               | X    | Y   | Z   | MO-Control<br><i>P</i> | MO-MI<br><i>P</i> |
|-------------------------------|-----|-----|-----|------------------------|-------------------|-------------------------------|------|-----|-----|------------------------|-------------------|
| Right hemisphere              |     |     |     |                        |                   | Cerebellar cortex             | -32  | -60 | -16 | —                      | *                 |
| Cerebellar cortex             | 06  | -82 | -20 | *                      | —                 | Cerebellar cortex             | -02  | -78 | -12 | *                      | *                 |
| Cerebellar cortex             | 08  | -82 | -16 | *                      | *                 | Lingual gyrus (a18)           | -02  | -76 | -08 | *                      | **                |
| Inferior temp. gyrus (a20)    | 56  | -32 | -16 | —                      | *                 | Middle occ. gyrus (a19)       | -42  | -66 | -08 | *                      | **                |
| Inferior temp. gyrus (a37)    | 46  | -52 | -16 | —                      | *                 | Substantia nigra              | -12  | -18 | -08 | *                      | —                 |
| Lingual gyrus (a18)           | 04  | -80 | -12 | *                      | *                 | Lingual gyrus (a18)           | -04  | -76 | -04 | **                     | **                |
| Fusiform gyrus (a18)          | 38  | -70 | -12 | —                      | *                 | Middle occ. gyrus (a19)       | -456 | -68 | -04 | **                     | **                |
| Inferior temp. gyrus (a20/21) | 48  | -30 | -12 | —                      | *                 | Lenticular nucleus            | -18  | 08  | -04 | *                      | —                 |
| Inferior temp. gyrus (a37)    | 48  | -54 | -12 | —                      | *                 | Lingual gyrus (a19)           | -06  | -60 | 0.0 | *                      | **                |
| Middle occ. gyrus (a19)       | 34  | -76 | -08 | *                      | **                | Inferior temp. gyrus (a37/19) | -46  | -66 | 0.0 | **                     | **                |
| Lingual gyrus (a18)           | 02  | -76 | -08 | *                      | **                | Inferior temp. gyrus (a37/19) | -46  | -68 | +4  | **                     | **                |
| Cingulate cortex (a32)        | 16  | 36  | -08 | *                      | —                 | Lingual gyrus (a18)           | -06  | -78 | +4  | —                      | **                |
| Cingulate cortex (a24)        | 18  | 34  | -04 | *                      | —                 | Middle occ. gyrus (a19/39)    | -44  | -70 | +8  | **                     | **                |
| Middle occ. gyrus (a19)       | 34  | -74 | -04 | *                      | **                | Lingual gyrus (a18)           | -10  | -82 | +8  | —                      | *                 |
| Inferior temp. gyrus (a37)    | 40  | -64 | 0.0 | *                      | **                | Cingulate cortex (a24)        | -20  | 34  | +8  | *                      | —                 |
| Lingual gyrus (a18)           | 02  | -76 | 0.0 | —                      | **                | Cingulate cortex (a24)        | -20  | 30  | +12 | *                      | —                 |
| Inferior temp. gyrus (a37)    | 50  | -56 | 0.0 | —                      | **                | Middle occ. gyrus (a19/39)    | -42  | -70 | +12 | *                      | **                |
| Inferior temp. gyrus (a37/21) | 36  | -66 | +4  | *                      | **                | Middle occ. gyrus (a19)       | -36  | -76 | +16 | —                      | *                 |
| Cingulate cortex (a24)        | 22  | 32  | +12 | *                      | —                 | Cingulate cortex (a24)        | -18  | 24  | +16 | *                      | —                 |
| Cingulate cortex (a24)        | 22  | 24  | +16 | *                      | —                 | Cingulate cortex (a24)        | -08  | 18  | +20 | *                      | —                 |
| Cingulate cortex (a24)        | 22  | 22  | +20 | *                      | —                 | Middle occ. gyrus (a19)       | -20  | -76 | +20 | —                      | **                |
| Cingulate cortex (a24)        | 10  | 24  | +24 | *                      | —                 | Cingulate cortex (a24)        | -20  | 10  | +24 | *                      | —                 |
| Left hemisphere               |     |     |     |                        |                   | Orbital gyrus (a19)           | -20  | -78 | +24 | —                      | **                |
| Cerebellar cortex             | -10 | -72 | -20 | *                      | *                 | Orbital gyrus (a19)           | -16  | -84 | +28 | —                      | *                 |
| Cerebellar cortex             | -02 | -78 | -16 | *                      | *                 |                               |      |     |     |                        |                   |

Coordinates are in mm, relative to the anterior commissure, corresponding to the Talairach and Tournoux atlas<sup>17</sup>. MO, movement observation; MI, motor imagery; a, Brodmann area; \*\*  $P < 0.01$  ( $z$  scores ranged from 2.3 to 3.0); \*  $P < 0.05$  ( $z$  scores ranged from 1.6 to 2.2); —, no variation. Seen from the rear of the brain, the X coordinate is horizontal (with positive values to the right), the Y

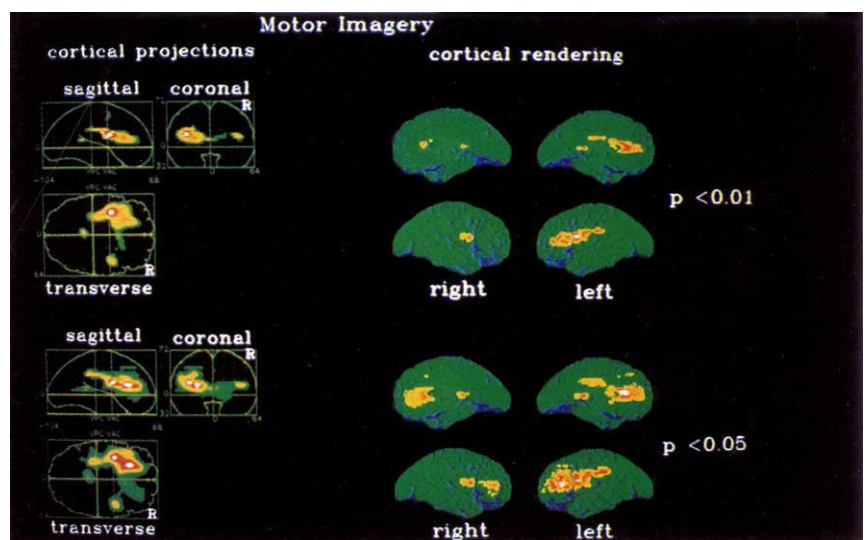
coordinate is in depth (with positive values anterior to the anterior commissure) and the Z coordinate is vertical (with positive values superior to the anterior commissure). After registration in the Talairach space, the analysis was carried out from Z coordinate -20 to +48 in all subjects, and from -20 to +52 in 5/6 subjects.

The pattern of activity during motor imagery was strikingly different. A significant activation was observed in areas concerned with motor behaviour. At the cortical level, area 6 in the inferior part of the frontal gyrus was strongly activated on both sides ( $P < 0.01$ ). On the left, contralateral to the side of the simulated movement, activation was also observed in prefrontal areas, extending to the dorsolateral frontal cortex (areas 9, 8

and 46), and in the inferior parietal lobule (area 40). Finally, the anterior cingulate cortex (areas 24 and 32) was bilaterally activated. At the subcortical level, the caudate nucleus was found to be activated on both sides ( $P < 0.01$ ), but this activation was more pronounced on the left. The cerebellum was involved only on the left side, and less than in the movement observation condition (Table 2 and Fig. 2).

FIG. 2 Adjusted mean rCBF while mentally imagining grasping objects with the right hand (MI) after subtraction of the visual inspection (control) condition. All conventions as in Fig. 1.

**METHODS.** The positron emission tomograph was a Siemens 931/04-12 (Siemens-CPS Knoxville, Tennessee, USA) whole-body scanner<sup>12</sup>. Three scans related to different experimental conditions were acquired for each subject, with a 20-min inter-scan interval. An i.v. bolus injection of 1,850 MBq  $H_2^{15}O$  was given for each run. Integrated counts were collected for 90 s starting 30 s after injection time. Two sets of 3 acquisitions each were done to cover the whole brain by moving the bed of the scanner in order to have an overlay on the last image section of the first set acquisition (bed = 0) with the first image section of the second set of acquisition (bed = 1). Once reconstructed, the images corresponding to the same condition (for example, motor imagery; bed = 0 and bed = 1) were combined to form a single image file. A normalization factor was calculated from the integral of counts detected in the overlaid plane section. Image transformation into a standard stereotactic anatomical space<sup>12</sup> was then done followed by statistical analysis using statistical parametric mapping. Global differences in CBF within and between subjects were covaried out for all voxels<sup>13</sup> and comparisons across conditions were made using the  $t$  statistic<sup>14</sup>. The significance of rCBF differences was assessed in an omnibus sense<sup>15,16</sup>. Threshold for significance was set at  $P < 0.05$  and  $P < 0.01$ . This assess-



ment of significance does not imply a correction for multiple comparisons and therefore few false positives may occur at the lower significance level ( $P < 0.05$ ). However, it is very unlikely that these explain the divergent patterns of activations observed during motor observation and motor imagery.



TABLE 2 Coordinates and *P* values for regions in which there was more activation during motor imagery

|                            | X   | Y   | Z   | MI-Control<br><i>P</i> | MI-MO<br><i>P</i> |                              | X   | Y   | Z   | MI-Control<br><i>P</i> | MI-MO<br><i>P</i> |
|----------------------------|-----|-----|-----|------------------------|-------------------|------------------------------|-----|-----|-----|------------------------|-------------------|
| Right hemisphere           |     |     |     |                        |                   | Sup. frontal gyrus (a10)     | -18 | 54  | +12 | —                      | *                 |
| Cingulate cortex (a24)     | 20  | 38  | -04 | *                      | —                 | Caudate nucleus              | -12 | 17  | +12 | **                     | *                 |
| Cingulate cortex (a24)     | 20  | 36  | 0.0 | *                      | —                 | Thalamus VL                  | -10 | 10  | +12 | —                      | *                 |
| Cingulate cortex (a24)     | 22  | 36  | +04 | *                      | —                 | Putamen                      | -30 | -10 | +12 | —                      | **                |
| Cingulate cortex (a24)     | 24  | 38  | +08 | *                      | —                 | Precentral gyrus (a6)        | -49 | 04  | +16 | **                     | **                |
| Sup. frontal gyrus (a10)   | 06  | 60  | +08 | —                      | *                 | Sup. frontal gyrus (a10)     | -20 | 50  | +16 | —                      | **                |
| Caudate nucleus            | 02  | 14  | +08 | *                      | *                 | Middle frontal gyrus (a46)   | -38 | 32  | +16 | —                      | **                |
| Sup. frontal gyrus (a10)   | 10  | 60  | +12 | —                      | *                 | Caudate nucleus              | -15 | 14  | +16 | *                      | **                |
| Precentral gyrus (a6)      | 42  | 08  | +12 | **                     | —                 | Sup. temporal gyrus (a22)    | -58 | -36 | +16 | —                      | **                |
| Caudate nucleus            | 04  | 16  | +12 | **                     | —                 | Sup. frontal gyrus (a10)     | -10 | 60  | +16 | —                      | *                 |
| Cingulate cortex (a24)     | 24  | 34  | +12 | *                      | —                 | Middle frontal gyrus (a46)   | -38 | 32  | +16 | —                      | *                 |
| Sup. frontal gyrus (a10)   | 16  | 58  | +16 | —                      | *                 | Cingulate cortex (a24)       | -22 | 24  | +16 | *                      | —                 |
| Precentral gyrus (a6)      | 42  | 06  | +16 | *                      | **                | Middle frontal gyrus (a46)   | -36 | 34  | +20 | —                      | *                 |
| Sup. frontal gyrus (a10)   | 18  | 56  | +20 | —                      | *                 | Sup. frontal gyrus (a10)     | -20 | 48  | +20 | —                      | *                 |
| Precentral gyrus (a6)      | 38  | 02  | +20 | *                      | *                 | Precentral gyrus (a6)        | -49 | 06  | +20 | *                      | *                 |
| Sup. frontal gyrus (a10)   | 18  | 52  | +24 | —                      | *                 | Precentral gyrus (a6)        | -34 | 00  | +24 | *                      | *                 |
| Cingulate cortex (a32)     | 02  | 32  | +24 | —                      | *                 | Middle frontal gyrus (a46)   | -44 | 24  | +24 | *                      | *                 |
| Precentral gyrus (a6)      | 28  | 02  | +24 | **                     | **                | Cingulate cortex (a23)       | -10 | -18 | +28 | **                     | —                 |
| Cingulate cortex (a32)     | 02  | 28  | +28 | —                      | *                 | Middle frontal gyrus (a46)   | -34 | 36  | +28 | —                      | *                 |
| Sup. frontal gyrus (a9)    | 20  | 46  | +28 | —                      | *                 | Precentral gyrus (a6)        | -30 | -08 | +28 | *                      | **                |
| Sup. frontal gyrus (a9)    | 20  | 40  | +28 | —                      | *                 | Middle frontal gyrus (a9)    | -44 | 24  | +28 | *                      | —                 |
| Precentral gyrus (a6)      | 30  | -02 | +28 | —                      | *                 | Inferior parietal lobule     | -34 | -28 | +28 | —                      | **                |
| Cingulate cortex (a32)     | 04  | 26  | +36 | —                      | **                | (a40)                        |     |     |     |                        |                   |
| Cingulate cortex (a32)     | 02  | 18  | +40 | *                      | **                | Postcentral gyrus (a1, 2, 3) | -26 | -22 | +32 | **                     | *                 |
| Cingulate cortex (a32)     | 04  | 10  | +44 | —                      | *                 | Middle frontal gyrus (a9)    | -44 | 22  | +32 | *                      | **                |
| Left hemisphere            |     |     |     |                        |                   | Inferior parietal lobule     | -28 | -30 | +32 | *                      | *                 |
| Cerebellar cortex          | -18 | -90 | -20 | *                      | —                 | (a40)                        |     |     |     |                        |                   |
| Caudate nucleus            | -06 | 20  | 0   | —                      | *                 | Cingulate cortex (a32)       | -04 | 22  | +36 | —                      | *                 |
| Caudate nucleus            | -06 | 10  | +04 | —                      | **                | Middle frontal gyrus (a9)    | -46 | 12  | +36 | *                      | **                |
| Sup. frontal gyrus (a10)   | -24 | 58  | +04 | —                      | *                 | Inferior parietal lobule     | -26 | -30 | +36 | *                      | *                 |
| Cingulate cortex (a24)     | -18 | 40  | +04 | *                      | —                 | (a40)                        |     |     |     |                        |                   |
| Sup. frontal gyrus (a10)   | -26 | 56  | +04 | —                      | *                 | Middle frontal gyrus (a9)    | -46 | 08  | +40 | —                      | *                 |
| Middle frontal gyrus (a46) | -36 | 46  | +08 | —                      | **                | Cingulate cortex (a32)       | -02 | 20  | +40 | —                      | *                 |
| Cingulate cortex (a24)     | -04 | 42  | +08 | *                      | —                 | Inferior parietal lobule     | -28 | -32 | +40 | *                      | *                 |
| Caudate nucleus            | -22 | 24  | +08 | **                     | **                | (a40)                        |     |     |     |                        |                   |
| Cingulate cortex (a24)     | -18 | 36  | +08 | **                     | —                 | Middle frontal gyrus (a6/8)  | -42 | 16  | +44 | *                      | *                 |

Coordinates are in mm, relative to the anterior commissure, corresponding to the Talairach and Tournoux atlas<sup>17</sup>. All notations and comments as in Table 1.

These results provide new insight into the mechanisms operating during the representation of motor actions. Activation of cortical areas and subcortical motor structures during both motor imagery and movement observation is a striking finding. Previous work with two-dimensional rCBF measurement had shown that the internal rehearsal of a motor sequence was associated with exclusive activation of the supplementary motor area<sup>6</sup>. In the present study, the ventral and mesial portions of the supplementary motor area were not found to be activated. Instead, more lateral premotor areas were involved. This can be explained by a difference in mental motor tasks: our task consisted of simulating visually guided movements, whereas the supplementary motor area exhibits preferential activity during internally guided tasks<sup>7</sup>. The involvement of prefrontal areas during motor imagery was only briefly mentioned<sup>8</sup>. Our results expand this finding in showing that consciously representing an

action involves a pattern of cortical and subcortical activation closely similar to that of an intentionally executed action. They further stress the role of dorsolateral frontal cortex in the generation of willed action, as already shown for movements actually performed<sup>9</sup>.

Finally, the activity in anterior cingulate cortex has been previously shown to be related to attention-requiring tasks<sup>10</sup>. By subtracting the two conditions from each other, we completely suppressed activation of area 24, whereas area 32 remained active in motor imagery. This finding confirms that part of anterior cingulate cortex plays an important role in higher motor control<sup>11</sup>. Our results altogether suggest that mental representations during observation of actions performed by others, and even more during simulation of one's own actions, share common neural mechanisms with other covert aspects of motor performance, such as planning and programming. □

Received 10 January; accepted 26 August 1994.

- Yue, G. & Cole, K. J. *J. Neurophysiol.* **67**, 1114–1123 (1992).
- Decety, J., Jeannerod, M., Germain, M. & Pastene, J. *Behav. Brain Res.* **42**, 1–5 (1991).
- Jeannerod, M. *Behav. Brain Sci.* **17**, 187–245 (1994).
- Watson, J. D. G. et al. *Cereb. Cortex* **3**, 79–94 (1993).
- Perrett, D. I. et al. *J. exp. Biol.* **148**, 87–113 (1989).
- Roland, P. E., Larsen, B., Lassen, N. A. & Skjold, E. *J. Neurophysiol.* **43**, 137–150 (1980).
- Mushiake, H., Inase, M. & Tanji, J. *J. Neurophysiol.* **66**, 705–718 (1991).
- Fox, P. T., Pardo, J. V., Petersen, S. E. & Raichle, M. E. *Neurosci. Abstr.* **14**, 1433 (1987).
- Frith, C. D., Friston, K., Liddle, P. F. & Frackowiak, R. S. J. *Proc. R. Soc.* **244**, 241–246 (1991).
- Pardo, J. V., Pardo, P. J., Janer, K. W. & Raichle, M. E. *Proc. natn. Acad. Sci. U.S.A.* **87**, 256–259 (1990).
- Paus, T., Petrides, M., Evans, A. C. & Meyer, E. *J. Neurophysiol.* **70**, 453–469 (1993).

- Spinks, T. J., Jones, T., Gilardi, M. C. & Heather, J. D. *IEEE Trans. nucl. Sci.* **35**, 721–725 (1988).
- Woods, R. P., Cherry, S. R. & Mazziotta, J. C. *J. Comp. Assist. Tomography* **16**, 620–633 (1992).
- Friston, K. J. et al. *J. cereb. Blood Flow Metab.* **9**, 690–695 (1989).
- Friston, K. J. *J. cereb. Blood Flow Metab.* **10**, 458–466 (1990).
- Friston, K. J., Frith, C. D., Liddle, P. F. & Frackowiak, R. S. J. *J. cereb. Blood Flow Metab.* **11**, 690–699 (1991).
- Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain* (Thieme, New York, 1988).

ACKNOWLEDGEMENTS. We thank M. Matarrese and G. Striano for their assistance in tracer synthesis and PET scanning and T. Locatelli for the electromyogram recordings. This work was supported by grants from the Human Frontier Science Program and from CNR. We also acknowledge support from the Ahmanson Foundation, the Pierson-Lovelace Foundation, the Brain Mapping Medical Research Organization and USPHS NINDS.

# Residual $\text{Ca}^{2+}$ and short-term synaptic plasticity

Haruyuki Kamiya\* & Robert S. Zucker†

Neurobiology Division, University of California, Berkeley, California 94720, USA

\* Present address: Department of Physiology, Faculty of Medicine, Kanazawa University, Kanazawa 920, Japan

† To whom correspondence should be addressed

At many synapses, the amount of transmitter released by action potentials increases progressively during a train of spikes. This enhancement of evoked transmitter release grows during tetanic stimulation with several time constants, each bearing a different name (facilitation: tens to hundreds of milliseconds; augmentation: several seconds; potentiation: several minutes), and the enhancement of release to test spikes after a tetanus decays with similar time constants. All these processes depend on presynaptic  $\text{Ca}^{2+}$  influx during the conditioning tetanus<sup>1</sup>. It has often been proposed that these forms of synaptic plasticity are due to residual  $\text{Ca}^{2+}$  present in nerve terminals following conditioning activity<sup>2</sup>. We tested this idea directly by using photolabile  $\text{Ca}^{2+}$  chelators to reduce residual  $\text{Ca}^{2+}$  following conditioning stimulation or to generate an artificial elevation in  $\text{Ca}^{2+}$  concentration, and observed the effects on synaptic transmission at crayfish neuromuscular junctions. We found that facilitation, augmentation and potentiation are caused by the continuing action of residual  $\text{Ca}^{2+}$ . Augmentation and potentiation seem to arise from  $\text{Ca}^{2+}$  acting at a separate site from facilitation, and these sites are different from the molecular target triggering neurosecretion.

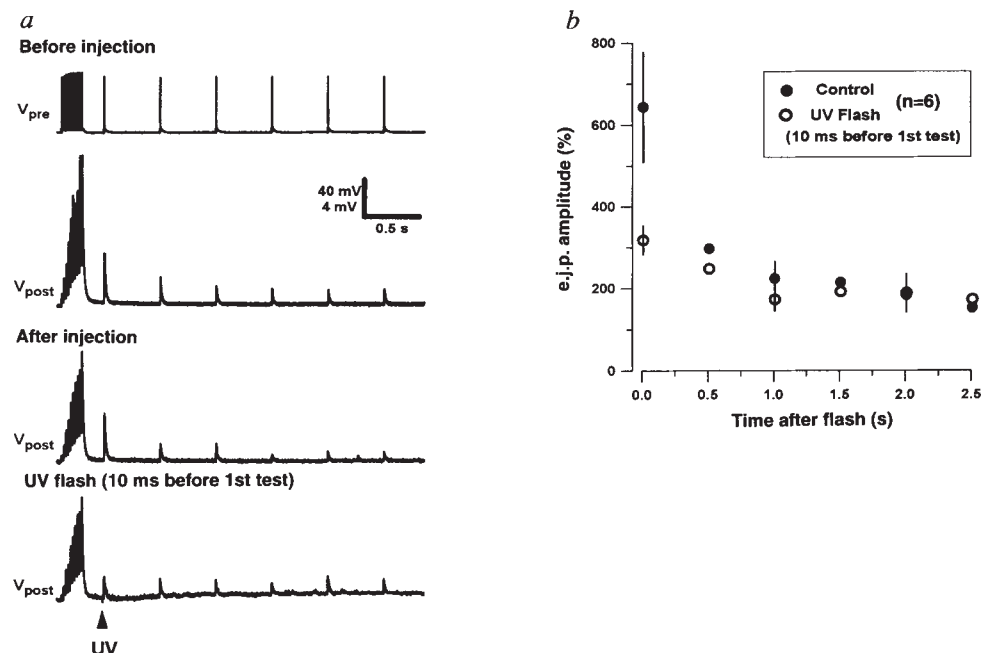
FIG. 1 Rapid reduction of residual  $[\text{Ca}^{2+}]_i$  quickly eliminates synaptic facilitation. **a**, Facilitation is established by 10 presynaptic stimuli at 50 Hz, and tested at 2 Hz beginning 200 ms after the end of the conditioning tetanus, when the first component of facilitation has decayed but the second component is still strong<sup>29</sup>. Top trace, action potentials recorded from a preterminal axon branch. Second trace, e.j.ps facilitate during the tetanus, and facilitation decays within 1.5 s. Third trace, after presynaptic injection of diazo-2, e.j.ps are slightly reduced. Bottom trace, a UV flash to photolyse diazo-2 occurs 10 ms before the first post-tetanic test response; production of presynaptic high-affinity  $\text{Ca}^{2+}$  buffer substantially eliminates facilitation. **b**, Rapid removal of facilitation by diazo-2 photolysis. Averaged e.j.p. amplitudes ( $\pm$ s.e.,  $n=6$ ) are expressed as percentages of unfacilitated e.j.ps before photolysis and plotted against time after flash. Similar results were obtained when the first post-tetanic e.j.p. followed the train by 10 ms as in these experiments or by 50 ms, indicating that both fast and slow components of facilitation are dependent upon residual  $\text{Ca}^{2+}$ .

**METHODS.** Dactyl opener muscles and motor nerves of isolated crayfish (*Procambarus clarkii*) walking legs were prepared as described previously<sup>12</sup>. The excitor motor neuron was penetrated near the first major Y branch with a bevelled 60–80 M $\Omega$  microelectrode containing 50 mM diazo-2, 10 mM fluorescein (both from Molecular Probes, Eugene, OR) and 100 mM K-HEPES, pH 7.3, and backfilled with 3 M KCl. E.j.ps were recorded at 14–16 °C with 4–10 M $\Omega$  microelectrodes filled with 3 M KCl from proximal-central fibres on the ventral surface near the presynaptic electrode. Diazo-2 was injected iontophoretically until

We used the photolabile  $\text{Ca}^{2+}$  chelator diazo-2 (ref. 3) to lower residual  $\text{Ca}^{2+}$  concentration ( $[\text{Ca}^{2+}]_i$ ) suddenly after conditioning stimulation and observed effects on short-term plasticity at excitor neuromuscular junctions in crayfish. Presynaptic injection of diazo-2 had little effect on the excitatory junctional potentials (e.j.ps) or their facilitation in a short train (Fig. 1a), as expected from the low buffer power of the unphotolysed chelator. Flash photolysis with ultraviolet light produces a high-affinity photoproduct which rapidly reduced facilitated e.j.ps tested 10 ms after photolysis nearly to the size of unfacilitated e.j.ps, while leaving unaffected the control e.j.ps recorded after facilitation that would normally have dissipated (Fig. 1b). The virtual abolition of facilitation within milliseconds of photolysis producing a step reduction in  $[\text{Ca}^{2+}]_i$  shows that facilitation is caused by residual  $\text{Ca}^{2+}$  acting at a site with fast kinetics which is in rapid equilibrium with residual  $\text{Ca}^{2+}$ .

We used a similar protocol to determine the role of residual  $\text{Ca}^{2+}$  in augmentation generated by a longer conditioning train and measured when facilitation had fully decayed. Figure 2a illustrates the post-tetanic decay of e.j.p. amplitudes tested twice per second. Photolysis of diazo-2 10 ms before an augmented e.j.p. had little immediate effect; later e.j.ps, however, were reduced nearly to control (non-augmented) levels (Fig. 2b). When photolysis preceded the test stimulus by 50 ms, augmentation was significantly reduced, and when the interval was set to 100 ms, the effect was even greater. The summary of these results (Fig. 2c) indicates that augmentation is blocked by reduction of residual  $[\text{Ca}^{2+}]_i$  with a time constant of  $\sim 350$  ms. A faster (50 ms time constant) component may also be present. Augmentation seems to be generated by residual  $\text{Ca}^{2+}$  acting at a slowly responding site different from that generating facilitation.

Next we looked at the role of residual  $\text{Ca}^{2+}$  in post-tetanic potentiation (PTP). Using an even longer conditioning train,



e.j.p. amplitudes just began to fall in muscles excited by terminals showing detectable fluorescein fluorescence; the final diazo-2 concentration was in the low millimolar range. A modified Chadwick-Helmuth (El Monte, CA) flashlamp was used to photolyse diazo-2, producing  $\sim 200$   $\mu\text{M}$  photoproduct with 150 nM  $\text{Ca}^{2+}$  affinity. This should have little effect on the  $[\text{Ca}^{2+}]_i$  peaks causing secretion<sup>25</sup>, but should readily absorb residual  $[\text{Ca}^{2+}]_i$  of a few  $\mu\text{M}$ <sup>14</sup>. In some experiments, diazo-4 (ref. 3) was used instead of diazo-2, with similar results. In control experiments, diazo-3 (ref. 3) was photolysed to produce photoproducts and a pH change similar to those of diazo-2 and diazo-4, but without a change in  $\text{Ca}^{2+}$  affinity; e.j.p. amplitudes were unaffected.



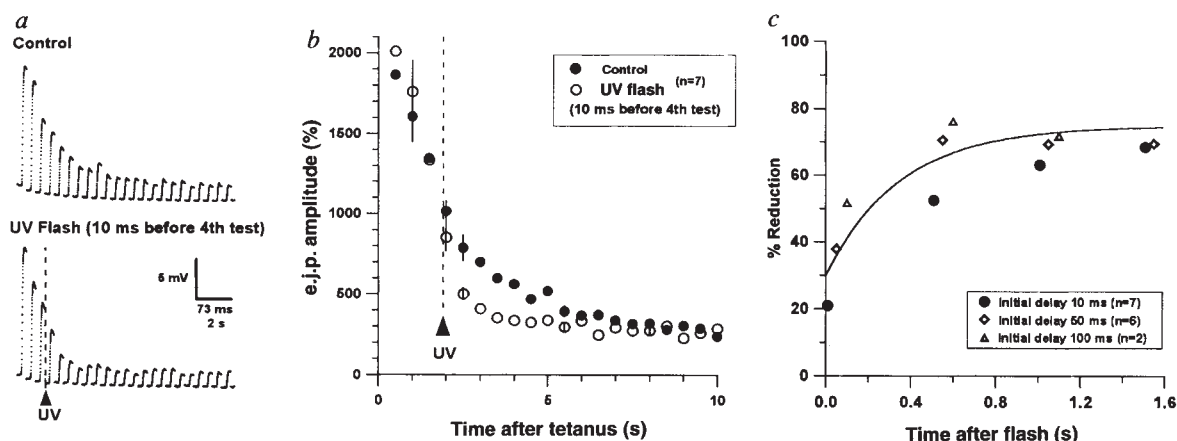


FIG. 2 Rapid reduction of residual  $[Ca^{2+}]_i$  eliminates synaptic augmentation after about 0.5 s. *a*, Augmentation is established by stimulating the motor neuron at 50 Hz for 4 s, and tested at 2 Hz beginning 500 ms after the end of the conditioning tetanus. Only postsynaptic responses are shown. Each response is displayed using a fast time scale (73 ms time calibration), but successive responses occur at 0.5 s intervals (2 s time calibration). In the lower trace, presynaptically injected diazo-2 is photolysed with a UV flash 10 ms before the fourth test response. This response is similar to the fourth response in the upper control trace without photolysis (but after diazo-2 injection); however, later responses

are markedly reduced after photolysis. *b*, Average reduction ( $n=7$ ) of augmentation by diazo-2 photolysis. e.j.p. amplitude is expressed as percentage of responses at 0.5 Hz. Photolysis causes little reduction of augmentation at 10 ms, but substantial reduction by 510 ms. *c*, Time course of block of augmentation expressed as percentage reduction of e.j.p. at different times after diazo-2 photolysis; 100% means that e.j.p. has been reduced to the final level 10 s after the tetanus, when augmentation is almost fully decayed. The flash preceded the first response by 10 ms, 50 ms, or 100 ms. The line is an exponential with time constant of 350 ms.

and waiting 1 min after the tetanus for facilitation and augmentation to decay fully, we observed the effect of diazo-2 photolysis on potentiated e.j.p.s. When the flash preceded an e.j.p. by 10 ms, little effect was observed, but PTP was nearly eliminated by the next e.j.p. 0.5 s later (Fig. 3*a*). The averaged results (Fig. 3*b, c*) show that PTP is blocked with kinetics similar to augmentation when residual  $Ca^{2+}$  is eliminated. As augmentation and potentiation are linearly correlated with similar sensitivities to micromolar levels of residual  $Ca^{2+}$  following presynaptic activity<sup>4,5</sup>, it is likely that these processes are caused by  $Ca^{2+}$  acting at one site with high affinity and slow kinetics, perhaps involving biochemical intermediaries.

Figures 2 and 3 show one difference between PTP and augmentation. Following diazo-2 photolysis, PTP always recovered to its previous level within ~30 s, whereas augmentation never did. Perhaps there is a recovery of residual  $Ca^{2+}$  after diazo-2 photolysis in PTP. This might occur if the intense long tetani required to establish PTP load intracellular  $Ca^{2+}$  stores, and slow leakage of  $Ca^{2+}$  from these organelles, as well as a reduction in  $Ca^{2+}$  removal by  $Na^+/Ca^{2+}$  exchange due to presynaptic  $Na^+$  accumulation<sup>6</sup>, produce the normal slow recovery of PTP. Leakage of  $Ca^{2+}$  from stores after diazo-2 photolysis might eventually saturate the small amount of photoproduct produced, leading to a rise in residual  $[Ca^{2+}]_i$  and the reappearance of PTP. We

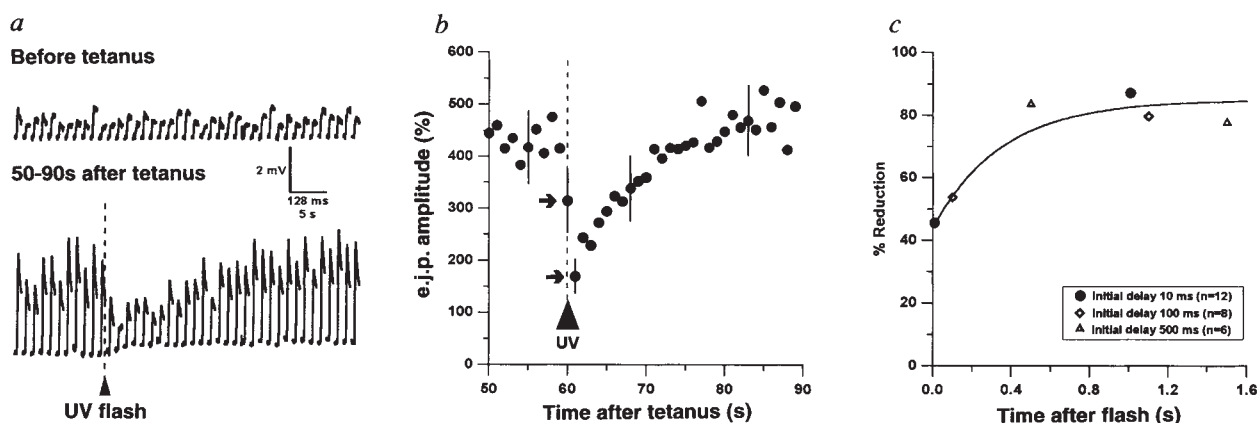


FIG. 3 After rapid reduction of residual  $[Ca^{2+}]_i$ , PTP dissipates at a similar rate as augmentation, but then recovers within 0.5 min. *a*, Potentiation is established by stimulating the motor neuron at 50 Hz for 5 min, and tested at 1 Hz after the tetanus. Forty postsynaptic responses are displayed in each trace; 25 ms of individual responses are displayed at fast sweep speeds as in Fig. 2. Upper trace, responses before the tetanus; lower trace, responses 50–90 s after the end of the tetanus, with a UV flash delivered 10 ms before the 61st post-tetanic response to photolyse diazo-2. At 1 min after the tetanus, facilitation and augmentation have decayed fully, but e.j.p.s are ~3 times their initial amplitude due to PTP. The first response after the flash is within the range of responses before photolysis, but the next few responses

are markedly reduced. PTP gradually recovers to its pre-existing level after photolysis. *b*, Average reduction ( $n=12$ ) of PTP by diazo-2 photolysis and subsequent recovery. e.j.p. amplitude is expressed as percentage of pre-tetanic control responses at 1 Hz. The arrows point to responses recorded 10 and 1,010 ms after the UV flash. *c*, Time course of block of PTP expressed as percentage reduction of e.j.p. at different times after diazo-2 photolysis. The flash preceded the first response by 10 ms, 100 ms, or 500 ms. The line is an exponential with time constant of 350 ms. As with augmentation, block of PTP by diazo-2 photolysis was not always complete, but the block shown at 10 ms intervals was not significantly different between PTP and augmentation (two-tailed *t* test,  $P>0.05$ ).

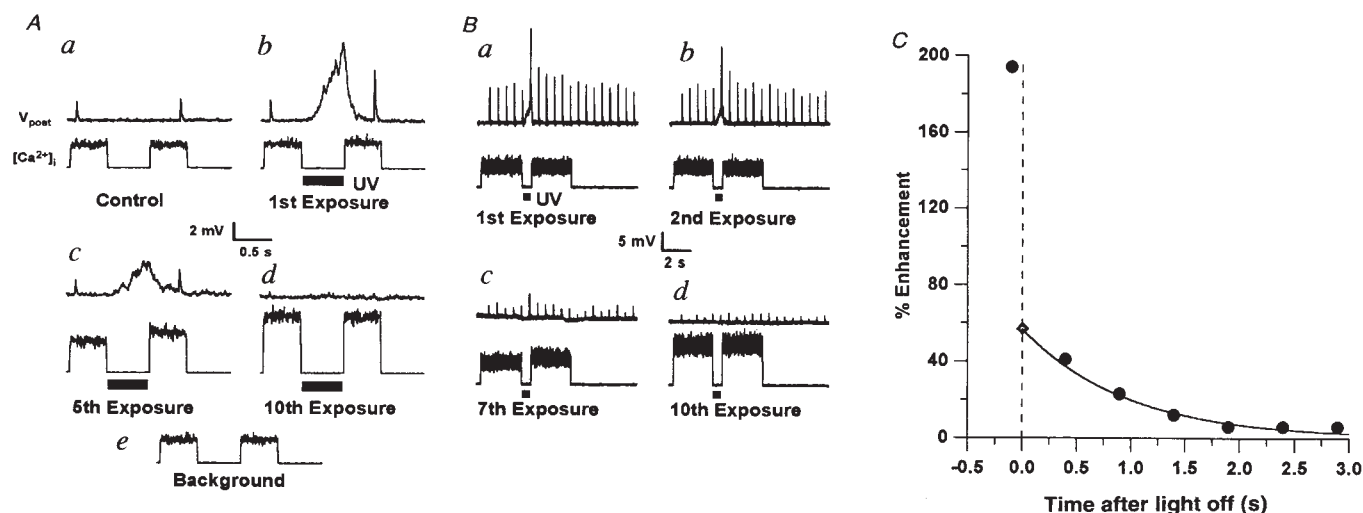


FIG. 4 Effect of reversible elevation of presynaptic  $[Ca^{2+}]_i$  on synaptic transmission. **A**, Steady photolysis of DM-nitrophen for 0.5 s elevates  $[Ca^{2+}]_i$  and evokes transmitter release during illumination (13 experiments). In each panel, top traces show postsynaptic potential; bottom traces monitor fluorescence of fluo-3 in terminals. In **a**, two 0.5-s  $[Ca^{2+}]_i$  measurements are made, indicated by a noisy departure of the  $[Ca^{2+}]_i$  trace from baseline. Evoked e.j.ps are relatively constant, and  $[Ca^{2+}]_i$  is low, so that fluo-3 fluorescence is close to background levels (shown in **e**). In **b**, the photolysis light is tuned on between the  $[Ca^{2+}]_i$  measurements: transmitter release occurs at an accelerating rate during light exposure, but ceases rapidly afterwards. The  $[Ca^{2+}]_i$  returns to low levels after photolysis, but e.j.ps are enhanced. By the 4th exposure  $[Ca^{2+}]_i$  has begun to rise after extinguishing the light, and rises more with the 5th (**c**) and successive exposures; transmitter release now persists at modest levels between illuminations. By the 10th exposure (**d**),  $[Ca^{2+}]_i$  has been in the  $\mu M$  range for over 1 min, and transmitter release is exhausted. In **B**, a similar experiment with 2 Hz stimulation shows the time course of the decay of enhanced e.j.ps after  $[Ca^{2+}]_i$  returns to a low level. In **C**, the decay in enhancement in e.j.ps amplitude (filled circles) is plotted against time following photolysis of DM-nitrophen for

0.5 s in a different preparation. This decay resembles that of augmentation and potentiation after diazo-2 photolysis. It is extrapolated to the e.j.p. enhancement that would be seen at the end of the photolysis period ( $\diamond$ ). This is substantially less than the enhancement shown by an e.j.p. during photolysis (first  $\bullet$ ). An additional facilitation-like process enhances spike-evoked release while  $[Ca^{2+}]_i$  is elevated, and it decays rapidly when photolysis stops. In six experiments, the extrapolated increase in e.j.p. amplitude after photolysis was  $52 \pm 10\%$  (mean  $\pm$  s.e.), whereas the increase in an e.j.p. measured 100 ms before stopping photolysis was significantly greater— $144 \pm 24\%$  ( $P < 0.05$ , two-sided *t*-test).

**METHODS.** The presynaptic electrode contained 45 mM DM-nitrophen (Calbiochem, San Diego), 13.5 mM  $CaCl_2$ , 0.9 mM Fluo-3 (Molecular Probes), and 57 mM KHEPES, pH 7.3. DM-nitrophen was injected by pressure to a final concentration in terminals of a few mM. Fluo-3 fluorescence was excited at 460–500 nm and monitored at 510–560 nm with a shuttered photomultiplier tube on the microscope's camera port. Photolysis was produced by a shuttered 150 W collimated Xe arc lamp. Calculations<sup>9,11,12</sup> suggest that  $[Ca^{2+}]_i$  increased to  $\sim 2 \mu M$  during light exposure.

tried to test this hypothesis by measuring  $[Ca^{2+}]_i$  in nerve terminals filled with diazo-2, but our efforts were thwarted by serious quenching of fluorescent indicators<sup>7</sup> by diazo-2.

If residual  $Ca^{2+}$  acts at low levels at separate sites to activate facilitation and augmentation/potentiation, elevation of  $Ca^{2+}$  to micromolar levels should elicit both forms of synaptic plasticity without conditioning stimulation. We injected nerve terminals with DM-nitrophen<sup>8</sup>, a photolabile high-affinity  $Ca^{2+}$  buffer which produces low-affinity photoproducts on photolysis. Exposure of DM-nitrophen partially loaded with  $Ca^{2+}$  to steady ultraviolet illumination causes  $[Ca^{2+}]_i$  to rise while the light remains on, and then return to near its initial level when the light is extinguished, so long as unphotolysed chelator remains in excess of total  $[Ca^{2+}]_i$  (ref. 9). Figure 4A shows the effect of 0.5 s illumination on presynaptic  $[Ca^{2+}]_i$  (measured with fluo-3, ref. 10), and postsynaptic potential. Asynchronous transmitter release increases during the light exposure to a rate of a few quanta per millisecond, causing several millivolts of postsynaptic depolarization<sup>11</sup>. This is intermediate between the resting miniature e.j.p. frequency (a few quanta  $s^{-1}$ ) and the maximal rate of secretion during an e.j.p. ( $\sim 250$  quanta  $ms^{-1}$ ). We cannot measure  $[Ca^{2+}]_i$  during exposure to the photolysing radiation, but calculate a rise to  $1\text{--}3 \mu M$ <sup>9,11,12</sup>. When illumination stops, secretion ceases and the fluorescence of fluo-3 is unchanged, indicating that  $[Ca^{2+}]_i$  has returned to a low level. Subsequent exposures evoke diminishing secretion, perhaps due to exhaustion of releasable transmitter stores or docked vesicles, or a side effect of DM-nitrophen photoproducts. By the fifth exposure, fluo-3 fluorescence has begun to increase and secretion continues after the light turns off, indicating that insufficient unphotolysed

DM-nitrophen remains to buffer the  $Ca^{2+}$  released by photolysis.

The reversible rise in  $[Ca^{2+}]_i$  increases e.j.ps evoked by action potentials (Fig. 4A). At 400 ms after the first exposure, when  $[Ca^{2+}]_i$  and transmission have returned to baseline levels, evoked transmission is still enhanced. Figure 4B shows that this enhancement decays with a time constant of several hundred milliseconds, similar to the decay of augmentation/potentiation after photolysis of diazo-2. This post-exposure enhancement of transmission is likely to be due to  $Ca^{2+}$  released from DM-nitrophen acting at the augmentation/potentiation site. The gradual rise in miniature e.j.p. frequency during illumination and a small tail of release lasting  $\sim 1$  s afterwards may also be due to  $Ca^{2+}$  acting at this site. Moreover, e.j.ps evoked during photolysis are increased even more than expected from the decay of the post-illumination enhancement (Fig. 4C). This effect, which decays rapidly on extinguishing the light, is probably due to  $Ca^{2+}$  released from DM-nitrophen acting at the site responsible for facilitation.

A popular hypothesis<sup>1,2</sup> attributes all forms of presynaptically enhanced transmission to the effect of a residual elevation in  $[Ca^{2+}]_i$ , summing with the change in  $[Ca^{2+}]_i$  accompanying action potentials to generate a higher peak  $[Ca^{2+}]_i$  transient acting on the molecular trigger for transmitter release. The non-linear dependence of transmitter release on  $[Ca^{2+}]_i$  (ref. 12) can generate a large enhancement of transmission in the face of little transmitter release due to residual  $Ca^{2+}$  itself. The different time constants of facilitation, augmentation and potentiation then reflect the complex time course of decay of residual  $Ca^{2+}$ . This idea is supported qualitatively by measurements of residual  $Ca^{2+}$  in nerve terminals associated with augmentation and



potentiation<sup>4,5,13</sup>. But recorded<sup>4,5,13</sup> or calculated<sup>14</sup> levels of residual  $[Ca^{2+}]_i$ , the rates of accumulation and decay of enhanced release<sup>2,15,16</sup>, and differential effects of  $Ba^{2+}$  and  $Sr^{2+}$  (refs 17, 18) do not accord quantitatively with predictions of the simple hypothesis of residual  $Ca^{2+}$  summing with  $Ca^{2+}$  influx to act at a single molecular target with nonlinear  $Ca^{2+}$  dependency.

Experimental tests of this hypothesis have focused on effects of presynaptically administered EGTA or BAPTA<sup>13,16,19–24</sup>; these tests have had variable results, sometimes blocking a component of enhancement of synaptic transmission, and sometimes not. Negative results may arise from exogenous chelator becoming saturated with  $Ca^{2+}$  during conditioning stimulation, whereas positive results could indicate that the exogenous chelator prevented  $Ca^{2+}$  from reaching sites causing synaptic plasticity that are distinct from the release trigger. Enhanced transmitter release could then be due either to actions of residual free  $Ca^{2+}$ , to  $Ca^{2+}$  remaining bound to sites of action, or to after-effects of  $Ca^{2+}$  that are independent of the continued presence of residual  $Ca^{2+}$ .

Our results show that facilitation, augmentation and potentiation are all due to the continuing action of residual free  $Ca^{2+}$  following neuronal activity. These processes are abolished by diazo-2 photolysis and activated by  $[Ca^{2+}]_i$  elevation. The similarity of the kinetics and calcium sensitivities of augmentation and potentiation suggest that these processes are due to  $Ca^{2+}$  acting at one site, with the prolonged duration of PTP reflecting the longer time it takes to remove residual  $Ca^{2+}$  after prolonged stimulation. The fast equilibration of the facilitation process with residual  $Ca^{2+}$  clearly distinguishes it from augmentation and potentiation, and suggests that  $Ca^{2+}$  acts at a separate site. It is also possible that facilitation arises from a more direct action of  $Ca^{2+}$  on some target, whereas augmentation and potentiation arise from subsequent reactions dependent on that same target. One question is whether facilitation can arise from the same low-affinity fast kinetic site as secretion<sup>25</sup>, as proposed by the original residual  $Ca^{2+}$  hypothesis. Attempts to simulate such a process have met with difficulty<sup>14</sup>, and led to the suggestion that facilitation is due to  $Ca^{2+}$  acting at a separate site from that causing secretion. Either this is a higher affinity site that is somewhat distant (at least a few hundred nanometres) from the local high peaks of  $Ca^{2+}$  very near  $Ca^{2+}$  channels, or its kinetics are not as fast as those causing secretion; otherwise it would be saturated by a single action potential. The recent observation of a delay in the development of facilitation at 0 °C (ref. 26) further distinguishes the facilitation site from the exocytosis site.

An important outstanding problem is the identification of the molecular targets of  $Ca^{2+}$  action in secretion and short-term plasticity. Many candidates exist among vesicular, plasma membrane, and cytoplasmic proteins, with little evidence to select between them<sup>27</sup>. A popular candidate for a target of short-term synaptic plasticity is synapsin I<sup>28</sup>, which can be phosphorylated by  $Ca^{2+}$ /calmodulin-dependent kinase II (CamKII). We measured the effects of presynaptic injection of calmodulin-binding domain (at 1 mM in the pipette with 10 mM fluorescein to monitor progress of injection), as well as bath application or presynaptic injection of the calmodulin inhibitor calmidazolium (up to 50  $\mu$ M in 0.5% DMSO in the bath or 0.2 mM in 2% DMSO in the pipette) and the CamKII inhibitor KN-62 (3  $\mu$ M added to 0.15% DMSO in the bath or 1 mM in 20% DMSO in the pipette); DMSO was added to the bath to prevent precipitation of hydrophobic substances in injected terminals. All treatments were without effect on transmission, facilitation, augmentation and potentiation. Such results do not favour calmodulin-mediated reactions as responsible for short-term synaptic plasticity. Identification of the molecular targets of  $Ca^{2+}$  awaits the development of more selective molecular probes. *Note added in proof:* A recent study of augmentation and PTP at hippocampal mossy fibre synapses<sup>30</sup> describes  $Ca^{2+}$ -dependent kinetics similar to, but somewhat slower than, those reported here for crayfish neuromuscular junctions. □

Received 14 April; accepted 8 August 1994.

1. Zucker, R. S. *A. Rev. Neurosci.* **12**, 13–31 (1989).
2. Magleby, K. L. & Zengel, J. E. *J. gen. Physiol.* **80**, 613–638 (1982).
3. Adams, S. R., Kao, J. P. Y. & Tsien, R. Y. *J. Am. Chem. Soc.* **111**, 7957–7968 (1989).
4. Delaney, K. R., Zucker, R. S. & Tank, D. W. *J. Neurosci.* **9**, 3558–3567 (1989).
5. Delaney, K. R. & Tank, D. W. *J. Neurosci.* (in the press).
6. Mulkey, R. M. & Zucker, R. S. *J. Neurosci.* **12**, 4327–4336 (1992).
7. Zucker, R. S. *Cell Calcium* **13**, 29–40 (1992).
8. Kaplan, J. H. & Ellis-Davies, G. C. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6571–6575 (1988).
9. Zucker, R. S. *Cell Calcium* **14**, 87–100 (1993).
10. Minta, A., Kao, J. P. Y. & Tsien, R. Y. *J. biol. Chem.* **264**, 8171–8178 (1989).
11. Mulkey, R. M. & Zucker, R. S. *J. Physiol. Lond.* **462**, 243–260 (1993).
12. Landó, L. & Zucker, R. S. *J. Neurophysiol.* **72**, 825–830 (1994).
13. Swandulla, D., Hans, M., Zipser, K. & Augustine, G. J. *Neuron* **7**, 915–926 (1991).
14. Yamada, W. M. & Zucker, R. S. *Biophys. J.* **61**, 671–682 (1992).
15. Bittner, G. D. & Schatz, R. A. *Brain Res.* **210**, 431–436 (1981).
16. Bain, A. I. & Quastel, D. M. *J. J. Physiol., Lond.* **455**, 383–405 (1992).
17. Zengel, J. E. & Magleby, K. L. *J. gen. Physiol.* **76**, 175–211 (1980).
18. Zengel, J. E. & Magleby, K. L. *J. gen. Physiol.* **77**, 503–529 (1981).
19. Tanabe, N. & Kijima, H. *Neurosci. Lett.* **99**, 147–152 (1989).
20. Delaney, K., Tank, D. W. & Zucker, R. S. *J. Neurosci.* **11**, 2631–2643 (1991).
21. Hochner, B., Parnas, H. & Parnas, I. *Neurosci. Lett.* **125**, 215–218 (1991).
22. Robitaille, R. & Charlton, M. P. *Annls N.Y. Acad. Sci.* **635**, 492–494 (1991).
23. Tanabe, N. & Kijima, H. *J. Physiol., Lond.* **455**, 271–289 (1992).
24. Van der Kloot, W. & Molgó, J. *J. Neurophysiol.* **69**, 717–729 (1993).
25. Adler, E. M., Augustine, G. J., Duffy, S. N. & Charlton, M. P. *J. Neurosci.* **11**, 1496–1507 (1991).
26. Van der Kloot, W. *J. Neurosci.* **14**, 5722–5724 (1994).
27. Bennett, M. K. & Scheller, R. H. *A. Rev. Biochem.* **63**, 63–100 (1994).
28. Linás, R., Gruner, J. A., Sugimori, M., McGuinness, T. L. & Greengard, P. *J. Physiol., Lond.* **436**, 257–282 (1991).
29. Zucker, R. S. *J. Physiol., Lond.* **241**, 91–110 (1974).
30. Regehr, W. G., Delaney, K. R. & Tank, D. W. *J. Neurosci.* **14**, 523–537 (1994).

ACKNOWLEDGEMENTS. We thank S. Adams and R. Tsien for a gift of diazo-4 and M. Melcher for confirming the activity of calmidazolium on yeast CamKII. This work was supported by a grant from the US NIH.

## Insulin-promoter-factor 1 is required for pancreas development in mice

Jörgen Jonsson, Lena Carlsson, Thomas Edlund\* & Helena Edlund

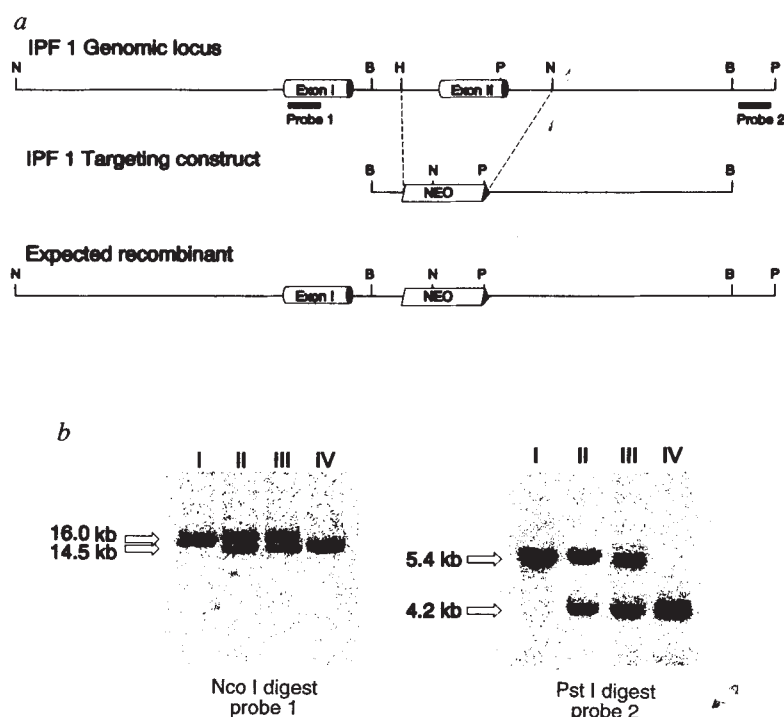
Department of Microbiology, University of Umeå, S-901 87 Umeå, Sweden

THE mammalian pancreas is a mixed exocrine and endocrine gland that, in most species, arises from ventral and dorsal buds which subsequently merge to form the pancreas. In both mouse and rat the first histological sign of morphogenesis of the dorsal pancreas is a dorsal evagination of the duodenum at the level of the liver at around the 22–25-somite stage, and shortly thereafter a ventral evagination appears as a derivative of the liver diverticulum<sup>1–3</sup>. Low levels of insulin gene transcripts are already present and restricted to the dorsal foregut endoderm at 20 somites, suggesting that pancreas- or insulin gene-specific transcriptional factors are present in this region before the onset of morphogenesis<sup>4</sup>. Insulin-promoter-factor 1 (IPF1) is a homeodomain protein which, in the adult mouse pancreas, is selectively expressed in the  $\beta$ -cells and binds to and transactivates the insulin promoter<sup>5</sup>. In mouse embryos, IPF1 expression is restricted to the developing pancreatic anlagen and is initiated when the foregut endoderm is committed to a pancreatic fate<sup>5</sup>. We now show that mice homozygous for a targeted mutation in the *Ipfl* gene selectively lack a pancreas. The mutant pups survive fetal development but die within a few days after birth. The gastrointestinal part and all other internal organs were normal in appearance. No pancreatic tissue and no ectopic expression of insulin or pancreatic amylase could be detected in mutant embryos and neonates. These findings show that IPF1 is needed for the formation of the pancreas and suggest that it acts to determine the fate of common pancreatic precursor cells and/or to regulate their propagation.

\* To whom correspondence should be addressed.

FIG. 1 Targeted disruption of the *Ipfl* gene. *a*, Schematic representation of targeting construct, genomic DNA and the expected product of homologous recombination. The two exons of *Ipfl* are indicated by cylinders and the bacterial neomycin gene, under control of the herpes simplex virus (HSV) promoter/enhancer, is represented by a triangular bar. Deletion of the 3.1-kb *HindIII*/*NcoI* fragment from within the 7.2-kb *BamHI* segment of the *IPF1* genomic DNA results in loss of the entire homeobox, the splice acceptor site and parts of the intron. This fragment was replaced with the 1,142-bp *XhoI*/*BamHI* fragment from pMC1 neopoly(A)<sup>12</sup>. Restriction enzymes: B, *BamHI*; H, *HindIII*; N, *NcoI*; P, *PstI*. *b*, Southern blot analysis of DNA from a wild-type sibling generated by a cross between two heterozygous *IPF1* mutants (I), the targeted E14 ES cell line (II) and hetero- (III) and homozygous *IPF1* mutants (IV). For confirmation of the correct recombination events at the 5' and 3' ends, genomic DNA isolated from ES-cell cultures or tail tips was digested with either *NcoI* and hybridized with probe 1, or with *PstI* and hybridized with probe 2, respectively. The genotypes (I–IV) are shown across the top, the restriction enzyme and the probe used are given below, and sizes of hybridizing fragments are indicated to the side of the blots.

**METHODS.** The mouse *Ipfl* gene was cloned from a mouse 129/SV genomic library using the mouse *Ipfl* cDNA clone as a probe<sup>5</sup>. E14-1 ES cells<sup>13</sup> were cultured on mitomycin-treated embryonic fibroblasts in medium supplemented with 1,000 U ml<sup>-1</sup> leukaemia-inhibitory factor as described<sup>13</sup>. A Bio-Rad GenePulser at 500 F, 260 mV was used to electroporate 10<sup>7</sup> cells with 25 µg ml<sup>-1</sup> linearized targeting DNA. Cells were plated on mitomycin-treated neomycin-resistant STO fibroblasts. Selection with 250 µg ml<sup>-1</sup> G418 was initiated after 48 h and resistant ES colonies were picked



eight days later. Blastocysts from C57BL/6 mice were injected and transferred to pseudopregnant (C57BL/6 × CBA)F<sub>1</sub> females to generate chimaeric offspring as described in ref. 14.

We previously isolated the cDNA that encodes the homeo-domain protein, insulin promoter factor 1 (IPF1)<sup>5</sup>. In early mouse embryos, the IPF1 protein is detected only in the developing pancreas but later in development and in adult mouse pancreas IPF1 is selectively expressed in the  $\beta$ -cells, where it binds to and transactivates the insulin gene<sup>5</sup>. The structurally related *Xenopus* X1Hbox-8 (ref. 6) and rat STF-1/IDX-1 (refs 7, 8) proteins are also selectively expressed in the endoderm of the duodenum and the pancreas but at present it is not known if these proteins represent functional homologues of IPF1. To test the hypothesis that IPF1 plays a role in the pancreatic commitment of the foregut endoderm, we generated *Ipfl*-deficient mice by deleting exon 2, which encodes the homeodomain of IPF1 (J.J. and H.E., unpublished results) using homologous recombination in ES-cells (Fig. 1). Mice heterozygous for the *Ipfl* mutation have no apparent abnormalities, they are fertile and their offspring show the expected mendelian frequencies of mutant genotypes, indicating that the *Ipfl*-deficiency does not cause embryonic lethality. However, all homozygous mutant mice die within a few days after birth, showing a complete penetrance of this neonatal mortality phenotype. The targeted *Ipfl*<sup>-/-</sup> mutant embryos show no detectable IPF1 immunoreactivity when analysed by whole-mount immunohistochemistry using anti-IPF1 antibodies (data not shown).

Newborn homozygous mutant mice do not show any morphological abnormalities, except that they appear slightly smaller than wild-type and heterozygous littermates: on average ~80% for newborn pups ( $n=15$ ) and ~60% for two-day-old pups ( $n=15$ ) (Fig. 2a). Most *Ipfl*-deficient pups are initially able to feed as they have milk in their stomachs, but all die within a few days after birth. To determine whether pancreas development was affected in *Ipfl* mutants, newborn pups from a cross between heterozygous *Ipfl* mutants were analysed histologically. Homozygous *Ipfl* mutants completely lack a pancreas but the duodenum from which the pancreas normally develops has the normal C-shaped form (Fig. 2b, c). The intestines of the *Ipfl*<sup>-/-</sup> pups ( $n=8$ ) have the same relative length (cm per g

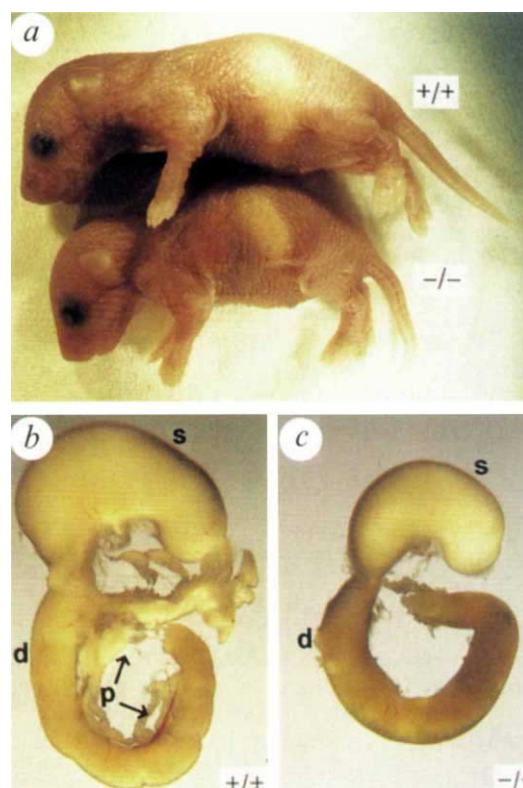


FIG. 2 *Ipfl*-deficient mice lack a pancreas. *a*, A pair of newborn littermates, wild-type (+/+) and *Ipfl* (-/-) mutant. *b*, Photograph of the stomach and duodenal region of a newborn wild-type mouse as it appears after removal from the abdomen. The arrows indicate the pancreas. *c*, Same as *b*, but from an *IPF1* (-/-) mouse. Note the complete absence of a pancreas. Abbreviations: s, stomach; d, duodenum; p, pancreas.



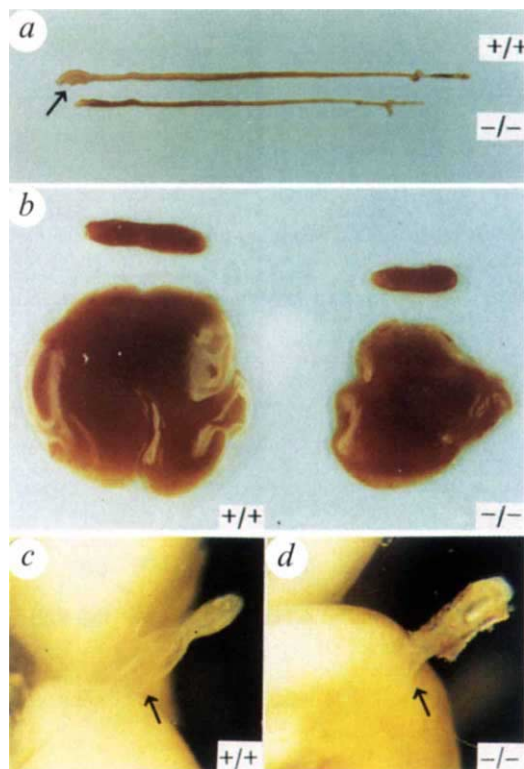


FIG. 3 The gut and its associated organs develop normally in *Ipfl* (-/-) mice. **a**, The intestines from a pair of 2-day-old with wild-type (+/+) and mutant (-/-) pups. Arrow indicates the pancreas. **b**, Spleens (top) and livers (bottom) from 2-day-old wild-type (left) and mutant (right) pups. **c**, The common bile duct of a newborn wild-type mouse. Arrow indicates the common bile duct. **d**, As **c**, but from an *IPF1* (-/-) mouse.

body weight),  $\pm 10\%$ , as the wild-type pups and show no apparent abnormalities except that the loops of the small intestine are positioned somewhat differently in the abdomen compared with the wild type (Fig. 3a, and data not shown). In the homozygous mutants ( $n=8$ ), both the liver, which develops from the same part of the primitive foregut as the pancreas, and the spleen, thought to be derived from 'pancreatic' mesoderm<sup>1</sup>, also appear to be normal and show the same relative weight (mg per g body weight),  $\pm 10\%$ , as wild type (Fig. 3b). The common bile duct and the ventral pancreatic duct are both derived from the hepatic diverticulum of the foregut and the main duct of the pancreas normally fuses with the common bile duct in the duodenal wall; both empty into the duodenal lumen at the major duodenal papilla<sup>9</sup>. In the homozygous mutants there is no pancreatic duct, but the common bile duct is present (Fig. 3c, d), indicating that, apart from the lack of a pancreas, the duodenal tract is normally developed. We conclude that *Ipfl* deficiency leads to the selective loss of the pancreas. *Ipfl*<sup>-/-</sup> pups that are able to feed and live for more than 2 days show elevated urine glucose levels,  $\geq 55$  mM for three-day-old *Ipfl*<sup>-/-</sup> pups ( $n=7$ ), suggesting that the cause of death is partly due to insulin deficiency. The lack of the other islet hormones and the exocrine digestive enzymes may also contribute to the pathology.

The complete lack of a pancreas indicates that IPF1 is required early in the development of the pancreas and suggests that IPF1 acts either at the level of determination or the early differentiation of the pancreas. In normal mice, pancreatic amylase and insulin are highly and specifically expressed in the exocrine and endocrine pancreas, respectively, and the expression of the gut-hormone gastrin can be used to determine the state of differentiation of the intestinal epithelium. To exclude the possibility that in *Ipfl*-deficient mice morphogenesis was arrested but cyto-differentiation still occurred, we analysed mutant and wild-type

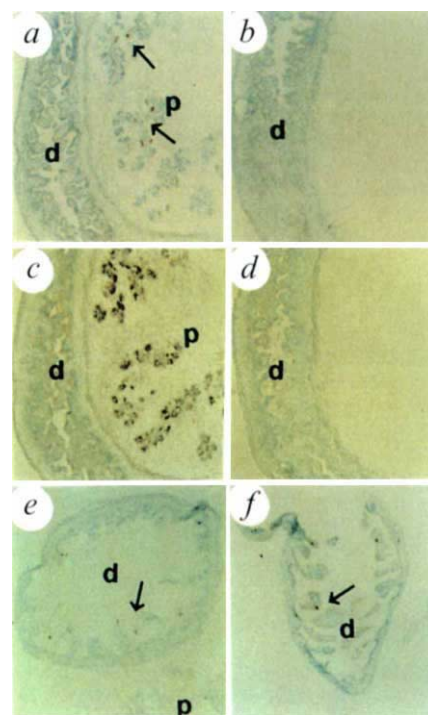


FIG. 4 *Ipfl* (-/-) embryos and neonates lack pancreatic tissue and show no ectopic expression of pancreatic exocrine and endocrine markers. High-power view of 7-μm sagittal (a-d) and transversal (e, f) cryostat sections of the duodenal-pancreatic tract from e15 wild-type embryos (a, c), e15 homozygous *IPF1* mutant embryos (b, d), and wild-type and mutant newborn animals (e, f), stained with anti-insulin antisera (a, b) anti-α-amylase antisera (c, d) and anti-gastrin antisera (e, f). Note the complete absence of a pancreas in the mutant embryos. Arrows in **a** indicate insulin-positive cells; in **e** and **f** they indicate gastrin-positive cells. Magnification,  $\times 28$ . Abbreviations: d, duodenum; p, pancreas. METHODS. Immunohistochemistry on mouse embryos and mouse tissue was done as before<sup>5</sup>, with the following modifications: embryos and tissues were fixed in 4% paraformaldehyde in 0.1 M sodium-phosphate buffer pH 7.4 for 1–2 h at 4 °C, washed in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% Triton X-100 and transferred to 30% sucrose in PB and incubated for 24 h at 4 °C before freezing and sectioning. Before immunostaining, sections were blocked for endogenous peroxidase activity in methanol containing 0.6% hydrogen peroxide for 20 min. IPF1 was detected with affinity-purified rabbit anti-IPF1 antibodies<sup>5</sup>, insulin with guinea-pig anti-human insulin antibodies (Linco), α-amylase with rabbit anti-human α-amylase antibodies (Sigma), and gastrin with rabbit anti-human gastrin antibodies (Dakopatt). The rabbit primary antibodies (anti-IPF1, anti-α-amylase and anti-gastrin) were detected with the ABC immunoperoxidase system (Vector Labs). Insulin antibodies were detected using peroxidase-conjugated rabbit anti-guinea-pig immunoglobulins (Dakopatt).

mouse embryos and neonates immunohistochemically. In the mouse there are enough insulin and amylase-expressing cells in the pancreas at around embryonic day 15 (e15) to be reproducibly detected by immunohistochemistry (refs 10, 11; Fig. 4a, c). The intestinal epithelium differentiates late in development, so expression of gastrin was monitored in sections of the duodenum in newborn animals. No pancreatic tissue was present in mutant e15 embryos and neonates (Fig. 4b, d, f) and no ectopic expression of insulin and amylase was detected in serial sagittal sections of the duodenum of mutant embryos (Fig. 4b, d) and neonates (data not shown). Cells expressing gastrin were present in the duodenum from both wild-type and mutant newborn animals (Fig. 4e, f). This, and the normal histology of the intestinal epithelium in the mutants indicate that this part of the duodenum develops normally. The lack of pancreatic tissue and

of ectopic expression of insulin and pancreatic amylase in the developing duodenum show that both cytodifferentiation and morphogenesis of the pancreas is arrested in the homozygous mutants.

The observed phenotype suggests that IPF1 has an early function in the initial stages of pancreas development. A few IPF1 positive cells can first be detected in the gut region at around the 10–12-somite stage (ref. 5, and our unpublished data), which is when the foregut endoderm commits to a pancreatic fate<sup>1</sup>. This and the lack of a pancreas in the *Ipfl*-deficient mutants strongly suggest that IPF1 functions in the determination and/or maintenance of the pancreatic identity of common precursor cells, or in the regulation of their propagation. □

Received 24 June; accepted 6 September 1994.

- Wessells, N. K. & Cohen, J. H. *Dev Biol.* **15**, 237–270 (1967).
- Spooner, B. S., Walther, B. T. & Rutter, W. J. *J. Cell Biol.* **47**, 235–246 (1970).
- Pictet, R. & Rutter, W. J. in *Handbook of Physiology* Section 7 (eds Steiner, D. F. & Frenkel, M.) 25–66 (Am. Physiol. Soc., Washington DC, 1972).
- Gittes, G. K. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **89**, 1128–1132 (1992).
- Ohlsson, H., Karlsson, K. & Edlund, T. *EMBO J.* **12**, 4251–4259 (1993).
- Wright, C. V. E., Schneegelsberg, P. & De Robertis, E. M. *Development* **104**, 787–794 (1988).
- Leonard, J. et al. *Molec. Endocr.* **7**, 1275–1283 (1993).
- Miller, C. P., McGehee, R. E. Jr. & Habener, J. F. *EMBO J.* **13**, 1145–1158 (1994).
- Githens, S. in *Human Gastrointestinal Development* (ed. Lebenthal, E.) 669–683 (Raven, New York, 1989).
- Herrera, P. L. et al. *Development* **113**, 1257–1265 (1991).
- Spooner, B. S., Cohen, H. I. & Faubion, J. *Dev Biol.* **61**, 119–130 (1977).
- Thomas, K. R. & Capocchi, M. R. *Cell* **51**, 503–512 (1987).
- Kühn, R., Rajewski, K. & Müller, W. *Science* **254**, 707–710 (1991).
- Hogan, B., Constantini, F. & Lacy, E. in *Manipulating the Mouse Embryo: A Laboratory Manual* 194–196 (Cold Spring Harbor Laboratory, New York, 1986).

ACKNOWLEDGEMENTS. We thank K. Rajewski for the E14.1 ES cell line; T. Jessell for critical reading of the manuscript; A. Cockshutt for proof reading; H. Semb, U. Dahl and members of our laboratories for suggestions; and H. Alstermark, K. Falk and U.-B. Backman for technical assistance. This work was supported by grants from the Swedish Medical Research Council and the Juvenile Diabetes Foundation, New York.

## A positive feedback loop coordinates growth and patterning in the vertebrate limb

Lee Niswander<sup>\*†</sup>, Susan Jeffrey<sup>†</sup>, Gail R. Martin<sup>\*</sup> & Cheryll Tickle<sup>‡</sup>

<sup>\*</sup> Department of Anatomy and Program in Developmental Biology, School of Medicine, University of California at San Francisco, San Francisco, California 94143-0452, USA

<sup>†</sup> Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA

<sup>‡</sup> Department of Anatomy and Developmental Biology, University College and Middlesex School of Medicine, Windeyer Building, Cleveland Street, London W1P 6DB, UK

Limb development depends on signals from the apical ectodermal ridge and underlying mesenchyme<sup>1,2</sup>. Fibroblast growth factor (FGF) can replace the ridge<sup>3,4</sup> and, because *Fgf4* RNA is localized to the mouse posterior ridge<sup>5</sup>, we proposed that FGF4 is the endogenous ridge signal<sup>3</sup>. Ridge signals control limb outgrowth and maintain the zone of polarizing activity (ZPA) at the limb posterior margin<sup>6</sup>, which is important in limb patterning: a ZPA graft to limb anterior mesenchyme causes cell specification and mirror-image duplications<sup>1,2</sup>. Sonic hedgehog (SHH<sup>7,8</sup>) has polarizing activity, and *Shh* RNA co-localizes with ZPA activity, suggesting SHH is the endogenous polarizing signal<sup>7</sup>. We have investigated the molecular regulation of *Fgf4* and *Shh* expression. We report here that *Fgf4* expression in the ridge can be regulated by *Shh*-expressing cells. Moreover, *Shh* expression in mesenchyme can be activated by FGF4 in combination with retinoic acid. Once induced, *Shh* expression can be maintained by FGF4 alone, thus establishing a positive feedback loop between ZPA and ridge.

Classical studies suggest interactions between ridge and ZPA<sup>1,2,9</sup>. In the mouse, *Fgf4* RNA is detected in the half of the ridge nearest the ZPA<sup>5</sup>, suggesting *Fgf4* expression might be induced by ZPA signals. To test this, we first needed to isolate a probe for chick *Fgf4*. Although the gene isolated encodes a protein highly homologous to those encoded by human and mouse *Fgf4* genes (Fig. 1a) and *Fgf6* genes<sup>10</sup> (data not shown), we propose it is the chick *Fgf4* homologue because its expression pattern (Fig. 1b, c, and data not shown) closely matches that of mouse *Fgf4* (including the posterior ridge)<sup>5</sup>, and not *Fgf6*<sup>11,12</sup>.

The anterior region of the chick wing bud was used as a model system for studying the effect of polarizing activity on *Fgf4* expression in the ridge. We applied a bead soaked in retinoic acid (RA-bead), which mimics the effects of a ZPA graft<sup>1</sup>, under the anterior ridge and assayed by *in situ* hybridization for *Fgf4* RNA in the anterior ridge, where it is not normally detected (Fig. 1b, c). In retinoic acid-treated limbs, *Fgf4* expression in the ridge extends further towards the anterior than in untreated contralateral limbs (ratio of distance from anterior limb margin at body wall to anterior limit of *Fgf4* expression on retinoic acid-treated versus untreated side was  $0.51 \pm 0.03$ ,  $n=10$ ; Figs 1d and 2a). We estimate retinoic acid-treatment increases the length of the *Fgf4* expression domain by about 50% (ratio of length of *Fgf4* expression domain in retinoic acid-treated versus untreated limb was  $1.55 \pm 0.15$ ;  $n=5$ ). This was not due to an overall increase in ridge length because the distribution of RNA of genes expressed throughout the ridge (*Bmp2*, *Bmp4* and *Msx1*; reviewed in ref. 1) was not altered (ratio of expression domain lengths in retinoic acid-treated versus contralateral limbs was  $1.04 \pm 0.06$ ,  $n=6$ ; see also ref. 9).

Injection of *Shh*-expressing cells or virus into wing anterior mesenchyme results in duplications similar to those seen with retinoic acid treatment<sup>7</sup>. To test whether SHH induces *Fgf4* expression, we grafted cells expressing *Shh* to the limb anterior margin. Within 24 h, *Fgf4* RNA was detected in the overlying anterior ridge ( $n=10$ ; Fig. 1e). Therefore, either retinoic acid or *Shh* expression in cells beneath the ridge can lead to *Fgf4* expression in the ridge.

Because retinoic acid induces *Shh* expression in the anterior mesenchyme of an intact limb bud<sup>7</sup>, and *Shh*-expressing cells induce *Fgf4* expression in the ridge, retinoic acid might induce *Fgf4* expression in the anterior ridge through SHH. However, after retinoic acid application, *Fgf4* RNA is readily detectable in the anterior ridge within 18 h (Fig. 2a), whereas *Shh* RNA is rarely detectable in mesenchyme until about 24 h (Fig. 2b, c and ref. 7). These unexpected results raise the possibility that instead of SHH regulating *Fgf4* expression, FGF4 might conversely play a role in regulating *Shh* expression. To test this, we applied a RA-bead to limb anterior mesenchyme following ridge removal. Under these conditions *Shh* expression was not induced (Fig. 2d). However, when both a RA-bead and FGF4-bead were applied, *Shh* expression was induced (Fig. 2e). Thus induction of *Shh* expression by retinoic acid requires a signal from the ridge and FGF4 can provide this signal. However, retinoic acid does not exert its effect on *Shh* simply by inducing FGF4 production in the ridge because when FGF4 alone is applied, *Shh* expression in limb anterior mesenchyme is not induced (Fig. 2f).

To determine whether FGF4 acts to maintain *Shh* expression once it is induced, we examined the effect of ridge removal and application of FGF4 on normal *Shh* expression in the ZPA. After ridge removal, *Shh* RNA decreased in posterior mesenchyme, and by 20 h was undetectable (Fig. 2d–f, and data not shown). However, FGF4 application to posterior mesenchyme maintains *Shh* RNA at a level comparable to that in the contralateral limb (Fig. 2g). FGF4 application to apical instead of posterior mesenchyme, does not maintain *Shh* expression (Fig. 2h), supporting the idea that FGF4 protein does not diffuse widely<sup>3</sup>.

A ridge signal is required not only for *Shh* induction (Fig. 2d–f) but also for ectopic activation of *Hoxd* gene expression



**METHODS.** Clones were obtained by screening stage 12–15<sup>20</sup> chick embryo cDNA and genomic libraries (kindly provided by D. Wilkinson, National Institute of Medical Research, London, and C. Tabin, Harvard Medical School, Boston, respectively) by low stringency hybridization with mouse *Fgf4* cDNA<sup>21</sup> and sequenced. For whole-mount *in situ* hybridization, embryos were processed<sup>22</sup> using digoxigenin-labelled chick *Fgf4* coding-region probe. In *d* and *e*, a RA-soaked (0.1 mg ml<sup>-1</sup>) bead<sup>23</sup> or pellet of chick embryo fibroblasts infected with *Shh* retrovirus<sup>7</sup> was inserted under the anterior ridge of stage 19–21 embryos. *Fgf4* expression was not induced in control experiments with DMSO-soaked beads or non-infected fibroblasts (data not shown).

FGF4 was applied, ectopic *Hoxd* gene expression was not induced (Fig. 3c) and limbs were truncated with a humerus that was thickened and often split ( $n=20$ ; Fig. 3g and ref. 3). SHH, like RA, can induce *Hoxd* expression in anterior mesenchyme of an intact limb bud<sup>7</sup>, but it remains to be determined whether this activation requires a FGF/ridge signal.

Our data suggest a model (Fig. 4) for the mechanism of retinoic acid action in respecification of limb anterior mesenchyme: (1) within 18 h of exposure to retinoic acid *Fgf4* expression is induced in the anterior ridge (Fig. 2a); (2) FGF4 produced by the ridge in combination with retinoic acid, induces *Shh* expression in anterior mesenchyme underlying the *Fgf4*-expressing ridge (inferred from the finding that FGF4 and retinoic acid are sufficient to induce *Shh* expression (Fig. 2e); (3) FGF4 maintains *Shh* expression (inferred from the finding that FGF4 maintains *Shh* expression in posterior mesenchyme (Fig. 2g); conversely, it seems likely that SHH also maintains *Fgf4* expression. Although such maintenance is not demonstrated here, *Shh*-expressing cells can induce *Fgf4* expression (Fig. 1e). We propose that these reciprocal interactions create a positive feedback loop between the two signalling centres, which no longer requires retinoic acid because the induced changes are irreversible about 16 h after treatment<sup>16</sup>. Indeed, after polarizing activity is established, FGF4 is sufficient to maintain *Shh* expression/polarizing activity in posterior mesenchyme (see also refs

**a**

|       |       |                                                                                         |                                                               |       |
|-------|-------|-----------------------------------------------------------------------------------------|---------------------------------------------------------------|-------|
| cFGF4 | (1)   |                                                                                         | M L S A A A L L P A L L L G L L W P G A V R G R P P P         | (27)  |
| mFGF4 | (1)   |                                                                                         | M A K R G P T T G T * * * R V * * A * V V A L * D * * T A M A | (31)  |
| hFGF4 | (1)   |                                                                                         | M S G P G T A * V * * * * V * * A * * A * W * G * * G A A A   | (31)  |
|       |       |                                                                                         |                                                               |       |
| cFGF4 | (28)  | G R L P P G P R Q - - - R R W D A A L F A R S V A R L P - - A E R R D                   |                                                               | (57)  |
| mFGF4 | (32)  | - - A * N * T * H A E L G H G * * G - * V * * * L * * * * V A * Q P P Q                 |                                                               | (64)  |
| hFGF4 | (32)  | P T A * N * T L E A E L E * * * E S - * V * L * L * * * * V A * Q P K E                 |                                                               | (66)  |
|       |       |                                                                                         |                                                               |       |
| cFGF4 | (58)  | A A - R D - - G D Y L L G Y K R L R R L Y C N V G I G F H I Q V L P D G                 |                                                               | (90)  |
| mFGF4 | (65)  | * * V * S G A * * * * * L * * * * * * * * * * * * * * L * * * * *                       |                                                               | (100) |
| hFGF4 | (67)  | * * V Q S G A * * * * * I * * * * * * * * * * * * * * L * A * * * *                     |                                                               | (102) |
|       |       |                                                                                         |                                                               |       |
| cFGF4 | (91)  | R I D G I H S E N R Y S L L E I S P V E R G V V S I F G V R S G L F V A                 |                                                               | (126) |
| mFGF4 | (101) | * * G * V * A D T * D * * * * L * * * * Q * * * * * * * * * * A * R F * * *             |                                                               | (136) |
| hFGF4 | (103) | * * G * A * A D T * D * * * * L * * * * * * * * * * * * * * A * R F * * *               |                                                               | (138) |
|       |       |                                                                                         |                                                               |       |
| cFGF4 | (127) | M N S K G K L Y G S Q T H V N D E C K F K E I L L P N N Y N A Y E S R I                 |                                                               | (162) |
| mFGF4 | (137) | * S * R * * * F * V - P P F T * * * * * * * * * * * * * * * * A Y A                     |                                                               | (171) |
| hFGF4 | (139) | * S * * * * * * * - P P F T * * * * * * * * * * * * * * * * * Y K                       |                                                               | (173) |
|       |       |                                                                                         |                                                               |       |
| cFGF4 | (163) | Y P G M Y I A L S K N G R T K K K G N K V S P T M T V T H F L P R I I .                 |                                                               | (196) |
| mFGF4 | (172) | * * * * F M * * * * * * * * * * * * * * * * R * * * * * K * * * * * * * L - .           |                                                               | (204) |
| hFGF4 | (174) | * * * * F * * * * * * * * * * * * * * * * K * * * * * R * * * * * K * * * * * * * L - . |                                                               | (206) |

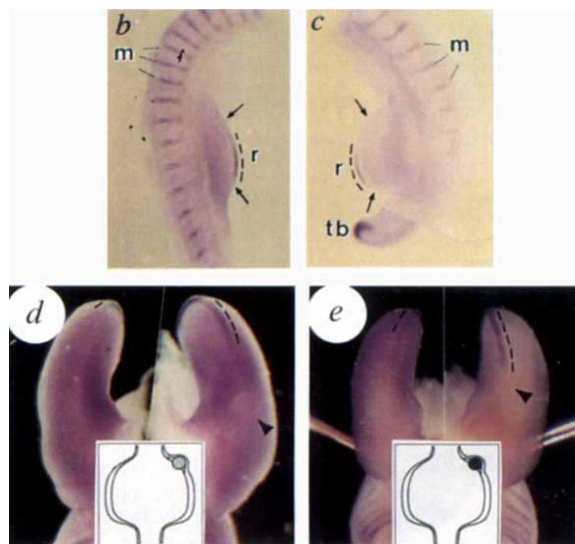


FIG. 2 Induction of *Shh* expression in limb mesenchyme in response to FGF4 and RA. Schematic drawings illustrate experimental manipulation; shaded circle represents RA-bead, open circle FGF4-bead. Arrowheads point to beads. a–c, Comparison of time at which *Fgf4* and *Shh* RNAs are detectable after RA-bead application under anterior ridge. RNA analysis using probes for a, *Fgf4* 18 h after RA-treatment; b, c, *Shh*, 19 and 24 h after RA-treatment, respectively. Dotted lines (a) mark *Fgf4* ridge expression, arrow (c) points to *Shh* expression in anterior mesenchyme (ectopic, RA-induced expression), expression is abundant in posterior limb mesenchyme (normal domain). d–h, *Shh* expression 24 h after ridge removal and application of (d) RA-bead to anterior mesenchyme; e, RA- and FGF4-beads to anterior mesenchyme (arrow indicates ectopic *Shh* expression); f, FGF4-bead to anterior mesenchyme; g, FGF4-bead to posterior mesenchyme; h, FGF4-bead to apical mesenchyme. Similar results were obtained 36 h after manipulation.

METHODS. RA-bead application (a–c) as in Fig. 1. In d–h, the ridge was removed and a RA-soaked ( $0.1 \text{ mg ml}^{-1}$ ) and/or FGF4-soaked ( $1 \text{ mg ml}^{-1}$  ( $\sim 50 \mu\text{M}$ ) recombinant human FGF4 protein kindly provided by Genetics Institute) bead was applied to exposed limb bud mesenchyme as previously described<sup>3</sup>. Eggs were incubated for time indicated and processed as in Fig 1 using chick *Shh*<sup>7</sup> or *Fgf4* probes.

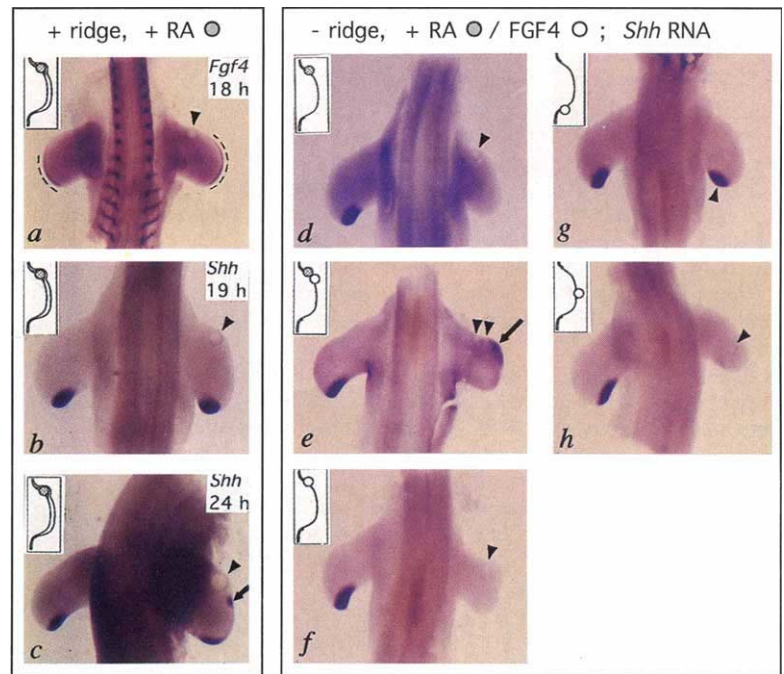


FIG. 3 *Hoxd-11* expression and skeletal formation after ridge removal and FGF4 and RA application a–c, Detection of *Hoxd-11* RNA<sup>15</sup> (purple stain) in chick wing buds at various times (25, 40 and 52 h) after ridge removal and application to anterior mesenchyme of a, RA- and FGF4-beads (40 h); b, RA-bead (40 h); c, FGF4-bead (25 h). Similar results were obtained at the different assay times. Unmanipulated limb on left. Arrowheads point to beads which sometimes absorb stain. Arrow (a) points to *Hoxd-11* RNA in anterior mesenchyme (ectopic, RA-induced expression; normal domain, posterior limb mesenchyme). d–g, Skeletal pattern 7 days after ridge removal and application of d, no bead (right limb), unmanipulated limb (left limb); e, RA- and FGF4-beads; f, RA-bead; g, FGF4-bead. In e, digits formed on anterior side (top) are posterior in character (digits 3 and 4) and anterior digit (digit 2) is not present as determined by the length of the metacarpals. Abbreviations: D, posterior digit-like elements; F, forearm bones; H, humerus; SG, shoulder girdle; 2, 3, 4, digits 2, 3, 4, respectively.

METHODS. Experimental manipulations and assays as in Figs 1 and 2. Skeletal elements visualized with alcian green<sup>3</sup>.

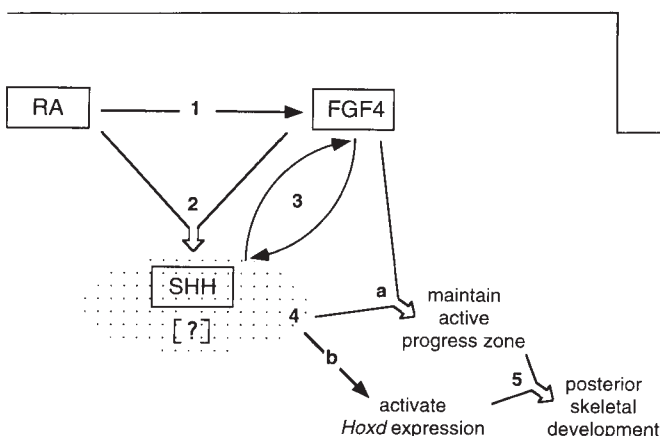
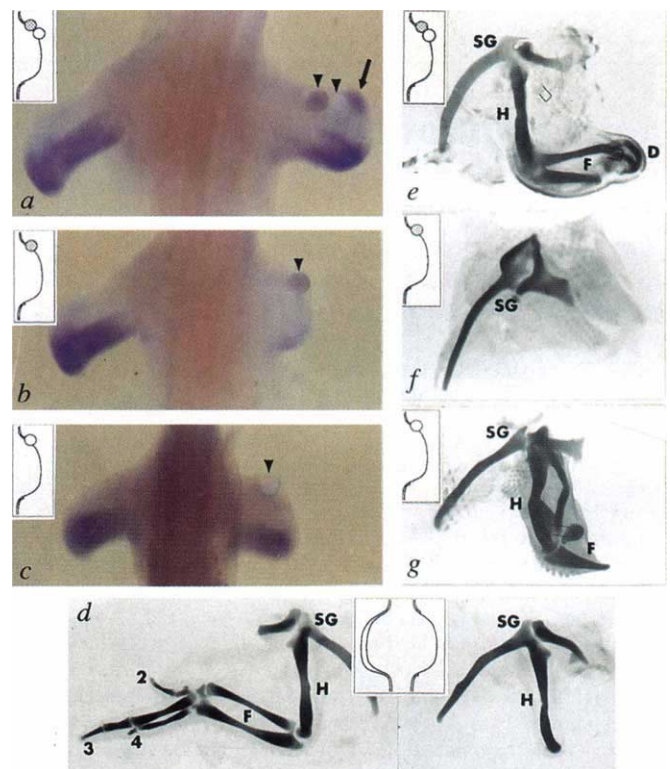


FIG. 4 Proposed model of molecular interactions induced in anterior ridge and mesenchyme by treatment with RA. (1) RA application under the anterior ridge induces *Fgf4* expression in the ridge. (2) FGF4, secreted by the ridge, and RA function together to activate *Shh* expression in the mesenchyme and establish a new ZPA (dotted area). (3) FGF4 and signals from new ZPA, presumably SHH itself, are mutually stimulatory and thus establish a positive feedback loop between ridge and mesenchyme signalling centres that is no longer dependent on RA. (4 and 5) Subsequent developmental events involved in respecification of anterior mesenchyme and formation of supernumerary skeletal elements. At each of these proposed steps the effect could be direct, or indirect and mediated by another gene or genes.



3, 6, 17). (4) The net result of expression of *Shh*, *Fgf4* and presumably other genes, is establishment and maintenance of a new ZPA in anterior mesenchyme. One proposed function of signal(s) from the ZPA, acting in conjunction with FGF4, is (i) to maintain an active progress zone<sup>3</sup>, where undifferentiated mesenchyme cells are assigned progressively more distal positional values, and (ii) to activate *Hoxd* expression and other downstream genes involved in mesoderm respecification. (5) This ultimately results in development of posterior skeletal elements from anterior mesenchyme. Although our data support the proposal that FGF4 is the key ridge-derived factor, other FGF family members expressed in the ridge may play a role, including FGF2<sup>4,18</sup>.

A positive feedback loop between the ridge and ZPA presumably plays a role in normal development, but how it is initially activated is unknown. *Shh* RNA is detected in posterior mesenchyme at about the same time as *Fgf4* RNA is detected in the posterior ridge (data not shown). This leaves open several possibilities, including induction of *Shh* expression by FGF4 in conjunction with endogenous retinoic acid present in the limb bud<sup>9</sup>, or induction of *Shh* expression in posterior mesenchyme by some other molecule(s), perhaps other FGF family members, leading to induction of *Fgf4* expression in the ridge. Once activated, the positive feedback loop serves as a simple mechanism whereby limb outgrowth and patterning are coordinately regulated. It is tempting to speculate that a similar regulatory loop involving the same gene families is used in other developmental processes, such as elaboration of the body axis. □

Received 11 August; accepted 31 August 1994.

1. Tickle, C. & Eichele, G. A. *Rev. Cell Biol.* **10**, 121–152 (1994).
2. Tabin, C. J. *Cell* **66**, 199–217 (1991).
3. Niswander, L., Tickle, C., Vogel, A., Booth, I. & Martin, G. R. *Cell* **75**, 579–587 (1993).
4. Fallon, J. F. et al. *Science* **264**, 104–107 (1994).
5. Niswander, L. & Martin, G. R. *Development* **114**, 755–768 (1992).
6. Vogel, A. & Tickle, C. *Development* **119**, 199–206 (1993).
7. Riddle, R. D., Johnson, R. L., Lauffer, E. & Tabin, C. *Cell* **75**, 1401–1416 (1993).
8. Roelink, H. et al. *Cell* **76**, 761–775 (1994).
9. Lee, J. & Tickle, C. *J. Embryol. exp. Morph.* **90**, 139–169 (1985).
10. Basilico, C. & Moscatelli, D. *Adv. Cancer Res.* **59**, 115–165 (1992).
11. Han, J.-K. & Martin, G. R. *Dev. Biol.* **158**, 549–554 (1993).
12. de Lapeyrière, O. et al. *Development* **118**, 601–611 (1993).
13. Koyama, E. et al. *Dev. Growth Differ.* **35**, 189–198 (1993).
14. Izpisua-Belmonte, J.-C., Brown, J. M., Duboule, D. & Tickle, C. *EMBO J.* **11**, 1451–1457 (1992).
15. Izpisua-Belmonte, J.-C., Tickle, C., Dollé, P., Wolpert, L. & Duboule, D. *Nature* **350**, 585–589 (1991).
16. Eichele, G., Tickle, C. & Alberts, B. M. *J. Cell Biol.* **101**, 1913–1920 (1985).
17. Anderson, R., Landry, M. & Muneoka, K. *Development* **117**, 1421–1433 (1993).
18. Dono, R. & Zeller, R. *Dev. Biol.* **163**, 316–330 (1994).
19. Thaller, C. & Eichele, G. *Nature* **327**, 625–628 (1987).
20. Hamburger, V. & Hamilton, H. L. *Dev. Dynam.* **195**, 231–272 (1992).
21. Hébert, J. M., Basilico, C., Goldfarb, M., Haub, O. & Martin, G. R. *Dev. Biol.* **138**, 454–463 (1990).
22. Izpisua-Belmonte, J.-C., De Robertis, E. M., Storey, K. G. & Stern, C. D. *Cell* **74**, 645–659 (1993).
23. Tickle, C., Lee, J. & Eichele, G. *Dev. Biol.* **109**, 82–95 (1985).

ACKNOWLEDGEMENTS. We thank our laboratory colleagues for helpful discussions, M. Tessier-Lavigne for critical reading of the manuscript, J. Zhang and C. Levine for library screening, C. Tabin for providing the *Shh* probe and retrovirus and D. Duboule for the *Hoxd* probes. This work was supported by the NIH Program of Excellence in Molecular Biology (G.R.M.), and MRC and Action Research grants (C.T.). L.N. was supported by a postdoctoral fellowship from the American Cancer Society, and by the Cancer Center Support Grant, Sloan-Kettering Institute.

## Activation of Raf-1 by 14-3-3 proteins

Wendy J. Fantl\*, Anthony J. Muslin\*, Akira Kikuchi, Jay A. Martin, Angus M. MacNicol\*, Richard W. Gross† & Lewis T. Williams\*

Daiichi Research Center, Cardiovascular Research Institute, Department of Medicine, University of California, San Francisco, California 94143, USA

† Division of Bioorganic Chemistry and Molecular Pharmacology, Washington University School of Medicine, St Louis, Missouri 63110, USA

**THE protein Raf-1, a key mediator of mitogenesis and differentiation, associates with p21<sup>ras</sup> (refs 1–3). However, the regulation of the serine/threonine kinase activity of Raf-1 is still not understood<sup>4–13</sup>. Using the yeast two-hybrid system<sup>8,14–16</sup>, we identified two structurally related proteins that interact with the amino-terminal region of Raf-1. These proteins, 14-3-3  $\zeta$  (PLA<sub>2</sub>) and 14-3-3  $\beta$  (HS1), are members of the 14-3-3 family of proteins<sup>17–23</sup>. Expression of 14-3-3 proteins in *Xenopus* oocytes enhanced Raf-1 activity and promoted Raf-1-dependent oocyte maturation. A dominant negative mutant of Raf-1 blocked the effects of 14-3-3 protein.**

Most of the complementary DNA clones identified in the two-hybrid screen<sup>14–16</sup> for Raf-1-interacting proteins encoded either 14-3-3  $\zeta$  or 14-3-3  $\beta$ <sup>17–21</sup> (see Table 1 legend). 14-3-3  $\zeta$  bound to the N terminal of Raf-1 roughly 40-fold better than to the C-terminal portion of Raf-1. 14-3-3  $\beta$  bound the N terminus of Raf-1 about 10-fold better than the C terminus (Table 1).

Myc-tagged 14-3-3 protein was expressed in *Xenopus* oocytes by RNA injection. Both 14-3-3  $\beta$  (HS1) and 14-3-3  $\zeta$  coimmuno-

TABLE 1 Detection of interactions between Raf-1 mutants and 14-3-3 proteins *in vivo* by two-hybrid screening

| Binding domain fusion | 14-3-3 $\zeta$<br>Units of $\beta$ -gal | HS1<br>Units of $\beta$ -gal |
|-----------------------|-----------------------------------------|------------------------------|
| c-Raf-1               | 4.8                                     | 6.2                          |
| 1098                  | 17                                      | 6.8                          |
| v-Raf                 | 0.4                                     | 0.9                          |
| Lamin                 | 0.3                                     | 0.3                          |

Yeast strain Y190 was cotransformed with the indicated panel of Raf-1 mutants and either 14-3-3  $\zeta$  or HS1. The efficacy of the interactions was quantified using the chlorophenylred- $\beta$ -D-galactopyranoside spectrophotometric assay<sup>29,20</sup>. Methods: Full-length human Raf-1, the N-terminal portion of Raf-1 (terminated after amino acid 323) or the C-terminal portion of Raf-1 (amino acids 1–320 deleted) were inserted into pAS1-CYH<sup>15</sup> as an in-frame fusion with the DNA-binding domain of GAL4 (residues 1–147) to make pAS1/Raf, pAS1/1098 or pAS1/vRaf. A HeLa cell two-hybrid cDNA library<sup>16</sup> that contained inserts fused to the transactivation domain of GAL4 (residues 768–881) was cotransfected into the yeast reporter strain Y190, which was then plated on media lacking histidine<sup>15,16</sup>. Of the 2.1 million yeast colonies expressing the two-hybrid constructs, 22 showed a specific interaction of a protein with Raf-1. Of these, 12 contained cDNAs encoding 14-3-3  $\beta$  and 2 clones encoded 14-3-3  $\zeta$ .

precipitated with Raf-1 in oocytes (Fig. 1a). When NAF (a point mutant of Raf-1 that can be recognized by Raf-1 antibodies, but has no kinase activity<sup>3</sup>) was coinjected with 14-3-3  $\beta$  the amount of 14-3-3 protein detected in a Raf-1 immunoprecipitate was increased. These findings show that 14-3-3 proteins and Raf-1 form a complex and that NAF can also interact with 14-3-3.

Microinjection of 14-3-3 either alone or with Raf-1 caused an increase in Raf-1 activity, assessed by measuring the ability of immunoprecipitated Raf-1 to activate MAP kinase kinase (Fig. 1b)<sup>24</sup>. The enhancement of Raf-1 activity by 14-3-3 was comparable to the increase in Raf-1 activity after progesterone treatment (Fig. 1b). NAF blocked activation of Raf-1 by 14-3-3, presumably by competing with Raf-1 for binding to 14-3-3.

Raf-1 activity regulates oocyte maturation<sup>25–27</sup>. Injection of 14-3-3 RNA (10–15 ng) with or without a Myc epitope tag, induced oocyte maturation as assessed by germinal vesicle break-

\* Present addresses: Cardiology Division, Department of Medicine, Jewish Hospital, Washington University Medical Center, St Louis, MO 63110, USA (A.J.M.); University of Chicago, 5841 South Maryland Avenue, Chicago, IL 60637, USA (A.M.M.); Chiron Corporation, Emeryville, CA 94608, USA (W.J.F., L.T.W.).

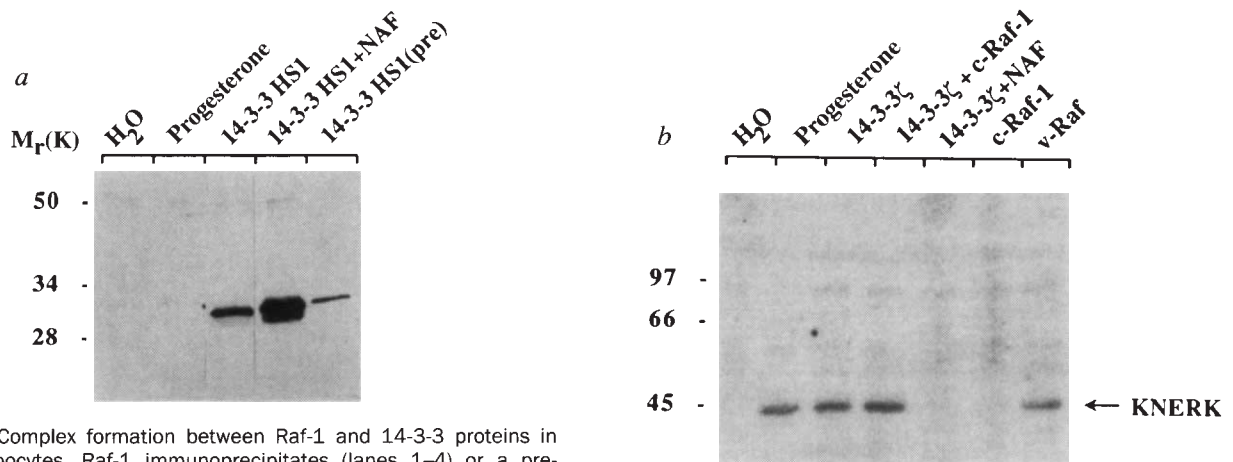


FIG. 1 a, Complex formation between Raf-1 and 14-3-3 proteins in *Xenopus* oocytes. Raf-1 immunoprecipitates (lanes 1–4) or a pre-immune precipitate (lane 5) were analysed by western blotting using an antibody to tagged 14-3-3 protein. RNA encoding 14-3-3 protein or NAF were injected as indicated in the last three lanes. Oocytes in the second lane were treated with progesterone.

METHODS. RNA encoding 14-3-3  $\beta$ , tagged at the N terminus with a Myc-derived epitope, was generated by *in vitro* transcription<sup>3,26</sup> and 10–15 ng of RNA was injected into oocytes. 24 h later, oocyte lysates were immunoprecipitated with Raf-1 antibody (Santa Cruz Biotechnology) or preimmune serum. A Myc antibody (Oncogene Science) was used for

western blots. b, Activation of Raf-1 by 14-3-3 proteins. The Raf-1 substrate GST-MEK and the substrate of MEK, kinase negative ERK (KNERK), were added to Raf-1 immunoprecipitates. An autoradiogram showing phosphorylation of KNERK is shown.

METHODS. Lysates of oocytes injected with the indicated RNAs (lanes 3–7) or treated with progesterone (lane 2) were immunoprecipitated with Raf-1 antibody and incubated with <sup>32</sup>P-ATP, GST-MEK and kinase negative ERK<sup>24</sup>.

down (GVBD). Western blotting verified the levels of 14-3-3 protein expression and confirmed that NAF did not alter these levels (data not shown). Conversely, 14-3-3 did not affect expression levels of Raf-1. In eight independent experiments, two examples of which are shown in Table 2, 25–60% GVBD was observed at 24 h in oocytes injected with 14-3-3, in contrast to 0% for controls injected with H<sub>2</sub>O, or with RNA encoding PDGF receptor (data not shown). Coinjection of c-Raf-1 did not enhance GVBD any further, whereas coinjection of NAF RNA (10–15 ng) with 14-3-3 RNA prevented GVBD.

Cdc2 kinase, a measure of maturation promoting activity<sup>28</sup>, was also activated by 14-3-3 protein to a level comparable to that induced by progesterone or v-Raf-1 (Fig. 2). Coinjection of NAF RNA with 14-3-3 prevented this activation.

In summary, we have cloned two 14-3-3 family members, HS1 and 14-3-3  $\zeta$ , that interact with the N-terminal part of Raf-1 in a yeast two-hybrid screen. A physical complex between 14-3-3

TABLE 2 Induction of oocyte maturation by expression of 14-3-3 protein

| RNA injected         | GVBD at 24 h/injected | GVBD(%) |
|----------------------|-----------------------|---------|
| Experiment 1         |                       |         |
| v-Raf                | 15/40                 | 38      |
| 14-3-3 $\zeta$       | 23/39                 | 59      |
| 14-3-3 $\zeta$ + NAF | 3/40                  | 7       |
| H <sub>2</sub> O     | 0/40                  | 0       |
| Progesterone-treated | 43/46                 | 93      |
| Experiment 2         |                       |         |
| v-Ras                | 18/53                 | 34      |
| myc-14-3-3 $\zeta$   | 11/39                 | 28      |
| myc-14-3-3 $\zeta$   | 0/38                  | 0       |
| myc-14-3-3HS1        | 11/38                 | 29      |
| myc-14-3-3HS1 + NAF  | 2/37                  | 5       |
| H <sub>2</sub> O     | 0/45                  | 0       |
| Progesterone-treated | 34/40                 | 85      |

Maturation is shown as the number of oocytes that have undergone germinal vesicle breakdown (GVBD) after injection with the indicated RNAs or after treatment with progesterone. Two representative sets of data are shown (experiments 1 and 2). Methods: Groups of 40 oocytes were scored for maturation by noting the appearance of a broad white spot that appeared at the animal pole of the oocyte<sup>26</sup>.

proteins and Raf-1 was also detected in oocytes, 14-3-3 proteins can dramatically activate Raf-1 and can stimulate the Raf-1-dependent process of maturation. Although it seems likely that the effect of 14-3-3 on Raf-1 activity is a direct consequence of the interaction of 14-3-3 with Raf-1, we cannot exclude the possibility that other molecules, such as p21<sup>ras</sup>, are involved in the effects of 14-3-3 proteins. It is now important to define the role of 14-3-3 proteins in a number of other Raf-1-dependent signal transduction processes. □

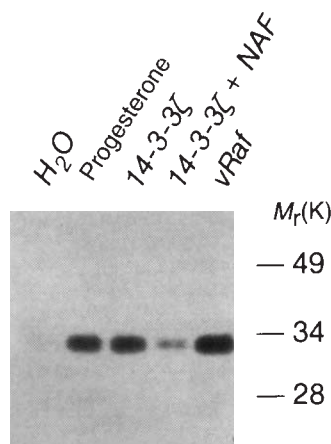


FIG. 2 Activation of Cdc2 kinase activity by 14-3-3 protein in oocytes. Cdc2 kinase was precipitated with GST-p13<sup>suc1</sup> agarose beads. An autoradiogram showing histone H1 phosphorylated by the Cdc2 kinase is shown.

METHODS. GST-p13<sup>suc1</sup> agarose beads (Upstate Biotechnology) were added to lysates of oocytes injected with the indicated RNAs (lanes 3–5) or treated with progesterone (lane 2). Histone H1 and <sup>32</sup>P-ATP were added to the precipitates as previously described<sup>28</sup>.

Received 1 August; accepted 19 September 1994.

1. Rapp, U. R. et al. *Oncogene* **6**, 495–500 (1991).
2. Kolch, W., Heidecker, G., Lloyd, P. & Rapp, U. R. *Nature* **349**, 426–428 (1991).
3. MacNicol, A. M., Muslin, A. J. & Williams, L. T. *Cell* **73**, 571–583 (1993).
4. Dent, P. et al. *Science* **257**, 1404–1407 (1992).
5. Howe, L. R. et al. *Cell* **71**, 335–342 (1992).
6. Kyriakis, J. M. et al. *Nature* **358**, 417–421 (1992).



**Wayne A. Fenton, Yechezkel Kashi\*,  
Krystyna Furtak† & Arthur L. Horwich‡‡**

Department of Genetics and † Howard Hughes Medical Institute,  
Yale University School of Medicine, Boyer Center,  
295 Congress Avenue, New Haven, Connecticut 06510, USA  
\* Present address: Department of Biotechnology, Technion Institute,  
Haifa 32000, Israel  
‡ To whom correspondence should be addressed

CHAPERONINS are ring-shaped protein complexes that are essential in the cell, mediating ATP-dependent polypeptide folding in a variety of compartments<sup>1-3</sup>. Recent studies suggest that they function through multiple rounds of binding and release of non-native proteins: with each round of ATP-driven release into the bulk solution, a substrate protein kinetically partitions between folding to the

native state or rebinding to another chaperonin molecule<sup>4-6</sup>. To gain further insight into the mechanism of polypeptide binding and release by the chaperonin GroEL from *Escherichia coli*, we have undertaken a mutational analysis that relates the functional properties of GroEL to its crystal structure<sup>7</sup>. Our functional tests identify a putative polypeptide-binding site on the inside surface of the apical domain, facing the central channel, consisting of hydrophobic residues. These same residues are essential for binding of the co-chaperonin GroES, which is required for productive polypeptide release. A highly conserved residue, Asp 87, positioned within a putative nucleotide-binding pocket in the top of the equatorial domain, is essential for ATP hydrolysis and polypeptide release.

Mutants of GroEL were devised mainly on the basis of examination of the three-dimensional model<sup>7</sup>. Residues within regions of interest that were selected for substitution are conserved either in all chaperonins—both Hsp60/GroEL and TF55/TCP1 families (Fig. 1, bold)—or only within the Hsp60/GroEL family (Fig. 1, orange). We overproduced mutant GroELs<sup>7</sup>, purified those present as assembled complexes, then determined their ATPase activity, GroES- and polypeptide-binding and polypeptide-folding abilities (Table 1). Their function *in vivo* was tested by the ability of the corresponding cloned mutant operons to rescue a

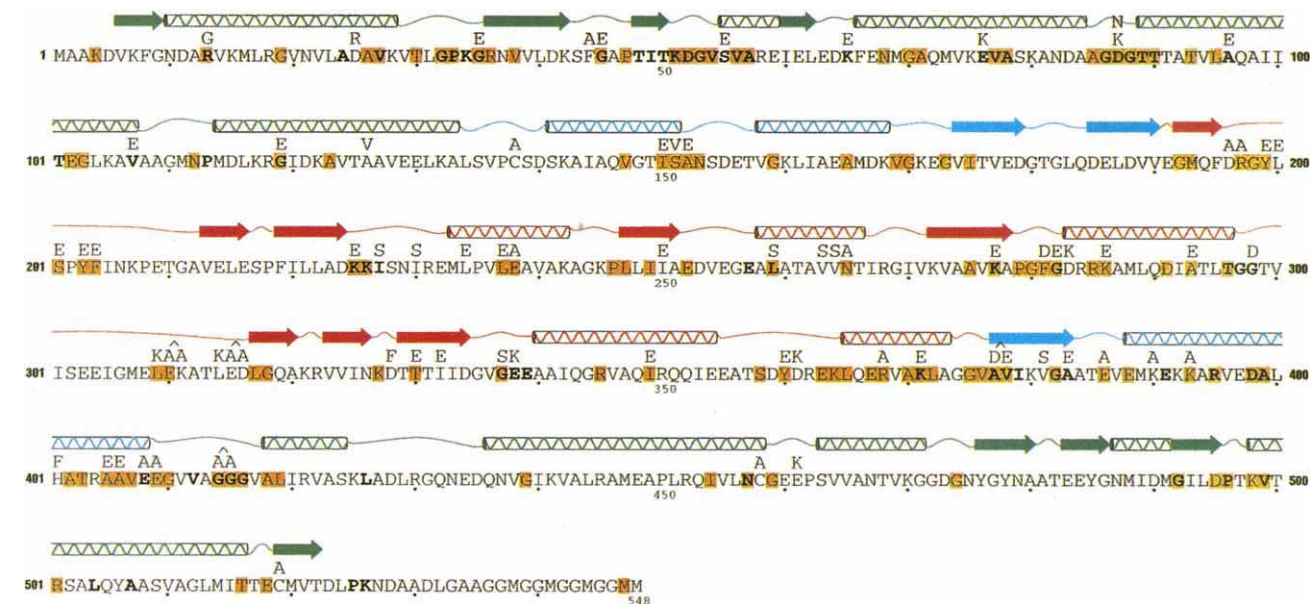


FIG. 1 Amino-acid substitutions in GroEL are shown above the wild-type amino-acid sequence. Secondary structure<sup>7</sup> (coil for  $\alpha$ -helix, arrow for  $\beta$ -sheet) is indicated by the overline, coloured according to domain: equatorial, green; intermediate, blue; apical, red. Residues conserved in both Hsp60/GroEL and TF55/TCP1 chaperonin families are shown in bold. Residues identical in at least 49 out of 50 compared Hsp60/GroEL sequences are shown on orange background.

by PCR using as template a wild-type *GroE* operon, and substituted for the wild-type sequence in a *trc*-expression plasmid. For each mutation, a restriction fragment bearing only the appropriate mutation as confirmed by sequence analysis, was recloned. All mutations are single-residue substitutions except for the following double mutants: R13G/A126V, G45E/I292T, E310A/K311A, E315A/D316A, G337S/I349E, A377D/V378E, E386A/K390A, K390A/K393A and G414A/G415A.

mutant GroEL-deficient strain<sup>8</sup> (LG6; Table 1). In general, those mutants found to be defective in one setting were also defective in the other.

Polypeptide binding was abolished by a number of mutations mapping to the inside face of the apical domain (Table 1 and Figs 2a, b, 3b). Binding was tested *in vitro* using a well established substrate, human ornithine transcarbamylase (OTC), a homotrimeric enzyme. This enzyme has an absolute requirement *in vivo* and *in vitro* for chaperonin-mediated folding<sup>8-10</sup>, becoming aggregated in the absence of chaperonin function. Productive release of OTC monomers from GroEL requires Mg-ATP and GroES and is followed by assembly into active trimer<sup>10</sup>. OTC subunits unfolded in guanidine hydrochloride were diluted into buffer containing wild-type or mutant GroEL, the mixture applied to a sucrose gradient, and the fractions analysed by SDS-PAGE and Coomassie-blue staining or western blotting.

When wild-type GroEL was in excess, virtually 100% of the OTC subunits were recovered with the complex (Fig. 2a). This was also observed with most of the mutant complexes (Table 1). By contrast, in the case of a number of mutations located in the apical domain (Y199E, Y203E, F204E, L234E, L237E, L259S, V263S and V264S) OTC was not associated with GroEL. In each case, OTC was detected as a precipitated aggregate at the bottom of the tube (Fig. 2a). The same mutant complexes also failed to bind another substrate, monomeric dihydrofolate reductase (DHFR) from chicken. When DHFR was diluted from denaturant into a solution containing a 2- or 3-fold molar excess of wild-type GroEL, DHFR was efficiently bound, preventing its spontaneous refolding to the active form (Fig. 2b)<sup>11</sup>. When the same experiment was performed with these apical mutants, DHFR promptly refolded in a manner identical to

TABLE 1 Functional analysis *in vitro* and *in vivo* of mutant GroEL molecules

| Mutation   | ATPase | GroES binding | Polypeptide binding | Folding | Strain LG6 | Mutation    | ATPase | GroES binding | Polypeptide binding | Folding | Strain LG6 |
|------------|--------|---------------|---------------------|---------|------------|-------------|--------|---------------|---------------------|---------|------------|
| R13G       | ↓      | +             | +                   | +       | +          | E238A       | +      | -/+*          | +                   | ↓/agg   | +          |
| R13G/A126V | ↓      | +             | +                   | +       | +          | L259S       | +      | -*            | -                   | -       | -          |
| D25R       | +      | +             | +                   | ↓       | small cols | V263S       | +      | -*            | -                   | -       | -          |
| F44A       | ↓      | +             | +                   | +       | -          | V264S       | ↓      | -*            | -                   | -       | -          |
| G45E       | +      | +             | +                   | ↓       | small cols | N265A       | +      | -*            | +                   | -/-rel  | -          |
| S55E       | +      | +             | +                   | +       | +          | K277E       | ↓      | +             | +                   | +       | small cols |
| K65E       | +      | +             | +                   | +       | +          | F281D       | ↓      | +             | +                   | ↓/agg   | -          |
| E76K       | +      | +             | +                   | +       | +          | D283K       | +      | +             | +                   | +       | +          |
| D87K       | -      | ↓*            | ↓                   | -/-rel  | -          | K286E       | +      | +             | +                   | +       | +          |
| D87N       | -      | ↓*/↑exch*     | +                   | -/-rel  | -          | A293E       | +      | +             | +                   | +       | +          |
| A96E       | +      | +             | +                   | +       | +          | L309K       | +      | +/↑exch*      | +                   | ↓/↓rel  | -          |
| V107E      | +      | +             | +                   | +       | +          | L314K       | +      | +/↑exch*      | +                   | ↓/↓rel  | +          |
| I150E      | -      | -*            | +                   | ↓/↓rel  | -          | G337S/I349E | +      | +/↑exch*      | +                   | ↓/agg   | -          |
| S151V      | -      | -*            | +                   | ↓/↓rel  | -          | E338K       | +      | +             | +                   | +       | +          |
| A152E      | -      | -*            | -                   | -       | tiny cols  | Y360E       | ↓      | +             | +                   | +       | -          |
| D196A      | +      | +             | +                   | +       | +          | D361K       | +      | -*            | +                   | ↓/↓rel  | +          |
| R197A      | +      | +             | +                   | +       | +          | R368A       | +      | +             | +                   | +       | +          |
| Y199E      | +      | -*            | -                   | -       | -          | K371E       | +      | +             | +                   | +       | +          |
| L200E      | +      | +             | +                   | +       | +          | V381S       | +      | +             | +                   | +       | +          |
| S201E      | +      | +             | ↓                   | ↓       | -          | A383E       | -      | -*            | +                   | -/-rel  | -          |
| Y203E      | +      | -*            | -                   | -       | -          | K390A/K393A | +      | +             | +                   | +       | +          |
| F204E      | +      | -*            | -                   | -       | tiny cols  | H401F       | +      | +             | +                   | +       | +          |
| K225E      | +      | +             | +                   | +       | small cols | A405E       | -      | +/↑exch*      | +                   | -/-rel  | -          |
| I227S      | +      | +             | +                   | +       | +          | A406E       | -      | -*            | ↓                   | ↓       | -          |
| I230S      | +      | +             | +                   | +       | +          | E409A       | +      | +             | +                   | +       | +          |
| L234E      | ↓      | -*            | -                   | -       | -          | E461K       | +      | +/↑exch*      | +                   | ↓/↓rel  | -          |
| L237E      | +      | -*            | -                   | -       | -          |             |        |               |                     |         |            |

Mutations producing the amino-acid substitutions shown in the first column were programmed into an overexpression plasmid bearing a *GroE* operon regulated by a *trc* promoter. The expressed mutant GroEL complexes were typically produced to a level ~100 times that of background wild-type complex (that is 50–75% soluble protein versus ~0.5%), so that the small background amount of wild-type subunits produced from the chromosomal operon would not be expected to interfere with the effect of mutation. Where intact 20S GroEL molecules were detected in cell extracts, they were further purified and assayed *in vitro* for the characteristics shown. ATPase activity was measured by determining free P<sub>i</sub> with malachite green. GroES binding was determined either indirectly by ability to inhibit the basal GroEL ATPase activity by ~50% (+ or -), or directly by ability of [<sup>35</sup>S]methionine-labelled GroES to cofractionate with GroEL in gel filtration (+\* or -\*). In a few cases the rate of exchange of [<sup>35</sup>S]GroES bound to GroEL with free, unlabelled GroES was tested and found to be increased compared with wild type (↑exch\*). Folding and release of a substrate polypeptide was estimated by the reconstitution of OTC activity from preformed OTC-GroEL binary complex. In some cases where folding was reduced (↓) or absent (-), the fate of the substrate polypeptide was determined: decreased release from GroEL (↓rel), no release (-rel), or release and aggregation (agg). For an *in vivo* test of function, *trc* plasmids bearing cloned mutant operons were transferred into the LG6 strain<sup>8</sup>, whose chromosomal *GroE* operon is regulated by a *lac* promoter. This strain is inviable in the absence of isopropylthiogalactoside (IPTG), but is rescued for growth in the absence of IPTG by transformation with a *trc* expression plasmid bearing a *GroE* operon with wild-type function<sup>8</sup>. ATPase activity of purified GroEL and its inhibition by GroES was measured essentially as described<sup>22</sup>. Briefly, about 20 µg GroEL was incubated at 37 °C for 5 min in 100 µl reconstitution buffer (Fig. 4 legend) with 1 mM ATP, with or without 0.4 µM GroES. The reaction was terminated with acidic Malachite green, sodium citrate was added and the absorbance at 650 nm recorded. Direct analysis of GroES binding is described in Fig. 2c legend. Exchange of radiolabelled GroES was measured in the presence of a 20-fold excess of unlabelled GroES and ADP by HPLC (as for Fig. 2c) after various intervals at room temperature. Typically, *t*<sub>1/2</sub> = 20 h for wild type; for affected mutants such as G337S/I349E, *t*<sub>1/2</sub> ≈ 2 min. Polypeptide binding was measured as described in legends to Fig. 2a and 4b; folding was estimated by the recovery of OTC activity in a reconstitution assay<sup>10</sup>; release was monitored by sucrose gradient analysis as in Fig. 4c. A number of mutants failed to form 20S molecules when overproduced, including G119E, G282E, A377D/V378E, E386A/K390A and E408A. The cloned operons in all cases failed to rescue LG6. Subunits of two mutants, G298D and D328F, were found to be cleaved in extracts of the expressing strain. Subunits of other mutants, T330 E and I332E, were insoluble. No mutant subunit was detected from strains containing plasmids I250E and G414A/G415A. Several mutants behaved as wild type in rescuing LG6, but were not fully analysed *in vitro*. These included G35E, G45S, G45E/I292T, C138A, E310A/K311A, E315A/D316A, C458A and C519A.



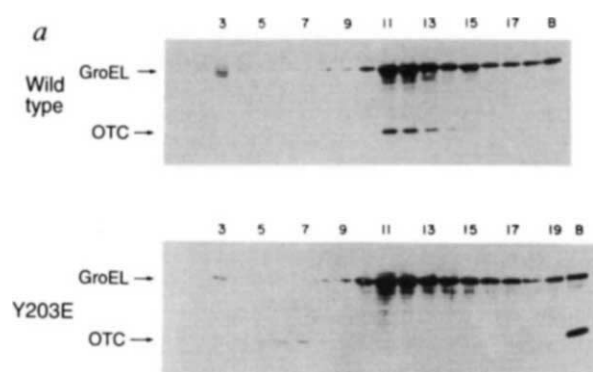
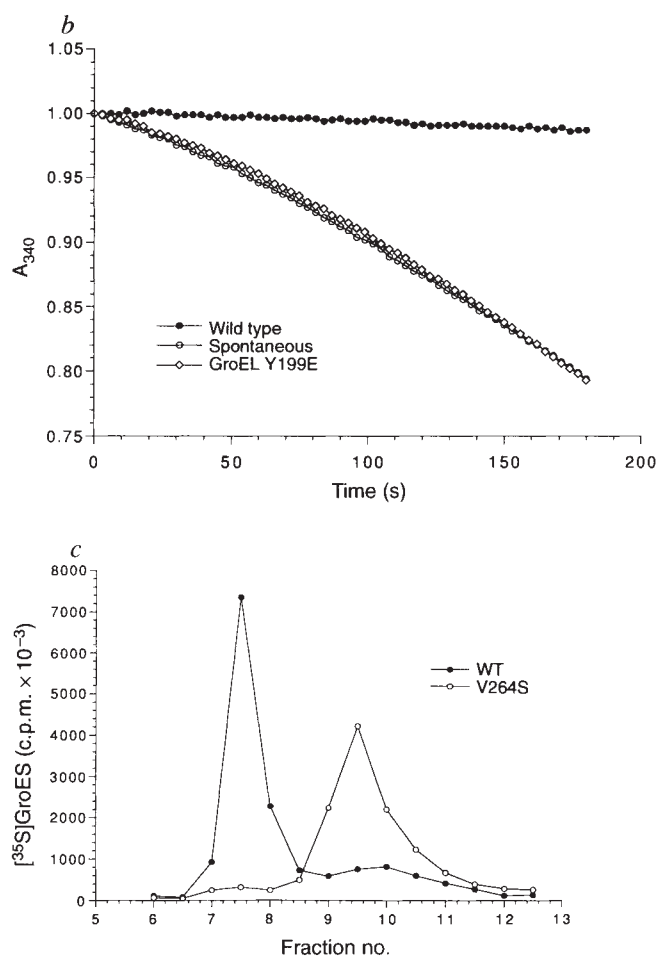


FIG. 2 Binding of substrate polypeptides and GroES is abolished by mutations in the apical domain. **a**, Binding of unfolded OTC to wild-type and Y203E mutant GroEL analysed by sucrose gradient fractionation. Unfolded OTC binds to wild type and migrates with GroEL, but fails to bind to Y203E, aggregates, and is recovered as a precipitate. **b**, Binding of unfolded DHFR to wild type and Y199E mutant GroEL determined by inhibition of its spontaneous refolding. Wild-type GroEL efficiently prevents the spontaneous refolding of unfolded DHFR, but mutant Y199E has no effect. **c**, Binding of native radiolabelled GroES to wild-type and V264S mutant GroEL detected by gel filtration chromatography. After incubation with wild-type GroEL, native GroES co-chromatographs as a complex with GroEL, but with the V264S mutant it chromatographs as an uncomplexed 70K species.

**METHODS.** **a**, Binding was carried out essentially as described<sup>10</sup>. Briefly, pure rat liver OTC was unfolded in 6M guanidinium hydrochloride (GuHCl) and diluted 50-fold into a solution containing equimolar (1 OTC monomer per GroEL 14-mer) purified wild type (top) or Y203E mutant (bottom) GroEL. After 30 min of incubation of 37 °C, the mixtures were cooled and applied to a 10–30% sucrose gradient in Beckman SW41 tubes. The gradients were centrifuged at 5 °C for 16 h at 36,000 r.p.m. ( $\omega^2 t = 8.1 \times 10^{11}$  s) in a SW41 rotor, then collected in 18 or 19 fractions. Aliquots of each fraction were treated with SDS sample buffer (5 min at 95 °C), then loaded onto an SDS–polyacrylamide (8%) gel and electrophoresed. The material precipitated in the bottom of the centrifuge tube (labelled 'B') was dissolved by adding SDS sample buffer directly to the empty tube and incubating briefly at 37 °C before recovering the material, heating and loading. After electrophoresis, gels were electroblotted onto a PVDF membrane (Immobilon-P) and reacted simultaneously with antisera against OTC and GroEL. Protein blots were developed with a chemiluminescent system (ECL, Amersham) and exposed to Hyperfilm-MP (Amersham). Only the relevant regions of the films are shown. The positions of GroEL and OTC were confirmed with standards and by reacting the blots independently with the two antisera. **b**, Spontaneous DHFR folding was inhibited essentially as described<sup>11</sup>. DHFR, denatured in GuHCl, was rapidly diluted 100-fold to 17 nM into an enzyme assay mixture containing 0.1 mM dihydrofolate and either no (spontaneous) or 140 nM wild type or Y199E mutant GroEL, followed immediately by addition of NADPH (0.1 mM) and mixing. The enzymatic reaction was followed by the decrease in absorbance (A) at 340 nm,



beginning ~10 s after NADPH addition. The initial absorbance was adjusted to 1.00 for each assay. The curvature of the spontaneous and GroEL Y199E plots reflects the increasing DHFR activity as a result of the continuing folding of the enzyme during the course of the assay. **c**, An *E. coli* strain overproducing GroES was grown in medium containing [<sup>35</sup>S]methionine, and the metabolically labelled [<sup>35</sup>S]GroES was purified by steps of anion-exchange chromatography. Specific activity of the purified radiolabelled GroES was ~50,000 c.p.m.  $\mu\text{g}^{-1}$ . Binding of GroES to GroEL was measured by incubating 640 nM GroEL (wild-type or Y199E mutant) with 160 nM [<sup>35</sup>S]GroES in a buffer containing 100 mM KCl, 50 mM Tris-HCl, pH 7.4, 5 mM MgCl<sub>2</sub> and 5 mM ADP, for 5 min at 20 °C. The mixture was analysed by HPLC on a Tosoh TSKGel-G4000SK column (0.78 × 30 cm) equilibrated in the same buffer, except with 1 mM ADP. Fractions of 1 ml were collected, and radiolabelled GroES counted in a scintillation counter.

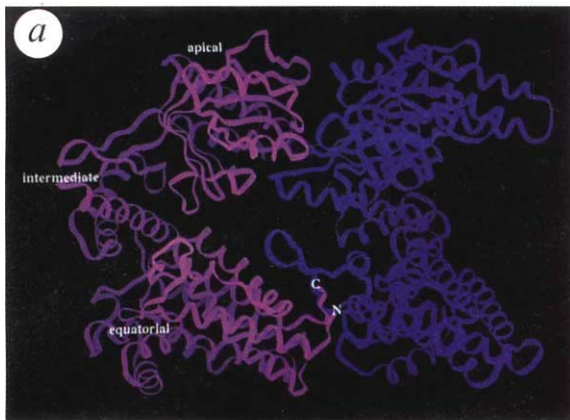
the spontaneous refolding observed in the absence of GroEL (Fig. 2b).

The binding-defective mutations map to the inside face of the apical domain, the least well resolved portion of the crystal structure<sup>7</sup> (Fig. 3b), which is most probably composed of two  $\alpha$ -helices and an underlying extended loop segment. This region lines the central cavity between the aperture and underside of the apical domain and could form a multivalent binding surface around the inside of the chaperonin cylinder. Notably, polypeptide binding was not affected by other substitutions elsewhere in the apical domain, including several mutations on the top surface of the domain in an extended loop segment (residues 298–315) and in an adjacent  $\alpha$ -helix (283–296) (Table 1 and Fig. 3). The role of the inside apical residues seems unlikely to be purely structural, considering that the region lies away from the central  $\beta$ -sheet structure of the apical domain and does not make contact with other regions of GroEL. Indeed, the mutant complexes did not exhibit any major rearrangement of architecture, as they

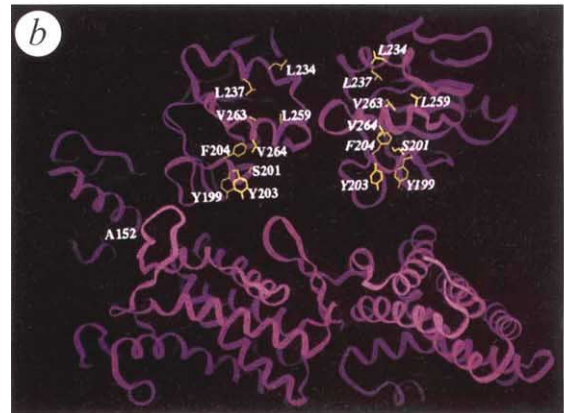
all had normal gel filtration behaviour and normal ATPase activity. Although the inside apical structures could conceivably gate polypeptide entry, we favour the idea that these residues provide a binding site(s) for non-native polypeptide. Such a function would be consistent with the poor resolution of this region in the crystal structure: the site could be inherently flexible to allow the surface to conform to a diverse range of unfolded polypeptides, enabling them to bind with high affinity.

Interestingly, the eight apical domain mutations that abolish polypeptide binding all involve residues that are hydrophobic in character (Fig. 3b). Notably, changes in charged residues within these regions, for example at residues 196, 197 and 238, were largely without effect (Table 1). These hydrophobic side chains in GroEL may therefore make contacts with hydrophobic sites in non-native folding intermediates<sup>12</sup>, sites that could otherwise dispose to aggregation. To determine how a non-native polypeptide is bound by the apical structure will ultimately require direct structural examination of a polypeptide–chaperonin complex,

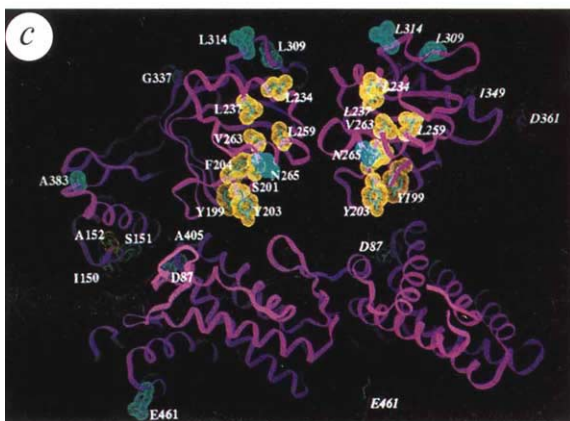
## Reference view of two subunits



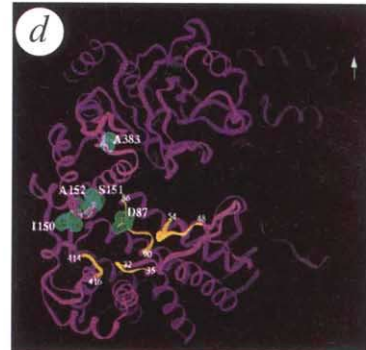
## Polypeptide binding



## GroES binding



## ATP hydrolysis



## Polypeptide folding

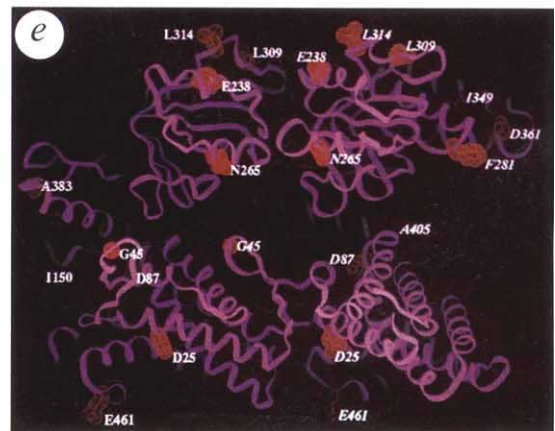


FIG. 3 Location of amino acids involved in GroEL functions: polypeptide binding, GroES binding, ATP hydrolysis and polypeptide release/folding. Each panel shows two neighbouring subunits of the 'top' GroEL ring, viewed from inside (from the position of the central channel). The ATP hydrolysis panel is an exception: the dimer has been horizontally rotated to view the ATP binding pocket of one of the subunits and the view is from the side, with outside on the left and the central channel to the right, with the axis of seven-fold symmetry designated by an arrow. *a*, Reference view of two subunits viewed from inside in full depth, with designation of domains and termini of one subunit. *b*, Residues involved in polypeptide binding form a site at the central face of the subunit apical domain. Note that this is the least resolved region of the GroEL structure (see Fig. 2c of the accompanying paper<sup>7</sup> for details of uncertainty). Side chains of the residues involved as positioned in the current model are indicated as yellow sticks. These residues, with exception of S201, have hydrophobic side chains. Mutation of residue 152 in the intermediate domain (a well resolved region) also interfered with binding, but this mutant is affected globally in that GroES binding and ATP hydrolysis are also defective (see text). *c*, GroES binding: the same residues that affect polypeptide binding are also required for binding of GroES, shown as yellow van der Waals spheres. Additional residues are also involved (blue), lying elsewhere in the apical domain (265, 309, 314, 337/349, 361), in the intermediate domain (150, 151, 152, 383), and in the equatorial domain at the ATP site (87) and the interface between rings (461). Note for 361 that the main chain site of attachment is not visible in this slab view (*a*, *e*). *d*, ATP hydrolysis: residues in which mutation ablated hydrolysis, shown as green spheres, map to the intermediate domain (150, 151, 152, 383, 406) and to the equatorial domain at the ATP-binding pocket (D87). Conserved residues around

the ATP-binding pocket are indicated by yellow ribbon and are designated with small numbers. *e*, Residues whose mutation produces defective release/folding of bound polypeptide. This group of mutants includes those shown to block GroES binding, which in general blocks polypeptide release, and those residues that block ATP hydrolysis. In addition, several mutants were defective in release and folding, despite normal GroES binding and ATP hydrolysis (at positions 25, 45, 238). Ribbon diagram was derived from the averaged model<sup>7</sup>, displayed in InsightII (BioSym Technology). The polypeptide main chain is shown as a purple ribbon; side chains are shown as sticks or sticks in Van der Waals spheres.

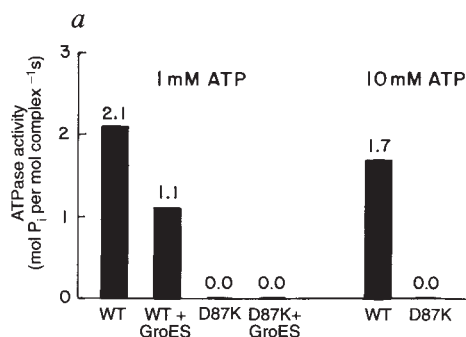
but the effects of the mutations studied here indicate that hydrophobic contacts between the non-native intermediate and chaperonin must contribute strongly. This idea is supported by the fact that the more conservative changes of Leu 237 to phenylalanine or alanine maintain binding competence

(results not shown), in contrast with the original change to glutamate.

Remarkably, the same apical mutations that block polypeptide binding also block GroES binding (Figs 2c and 3c). Binding was measured by incubating wild-type or mutant GroEL with



ADP and radiolabelled GroES, followed by gel-filtration analysis. Whereas radiolabelled GroES comigrated with wild-type GroEL, it failed to associate with the mutant complexes (Fig. 2c). The finding that GroES binding is abrogated in the same mutants defective in polypeptide binding suggests that polypeptide and GroES might compete for a common site of recognition (Fig. 3). That GroES physically interacts with this site is suggested by cryo-electron microscopy of GroEL–GroES complexes, in which the apical domains of GroEL appear to be opened outwards to make contact at their inner face with GroES<sup>13</sup>. With seven-fold symmetry fitting precisely to GroEL, GroES could act, perhaps through the previously identified mobile segment<sup>14</sup>, to displace polypeptide by direct competition.

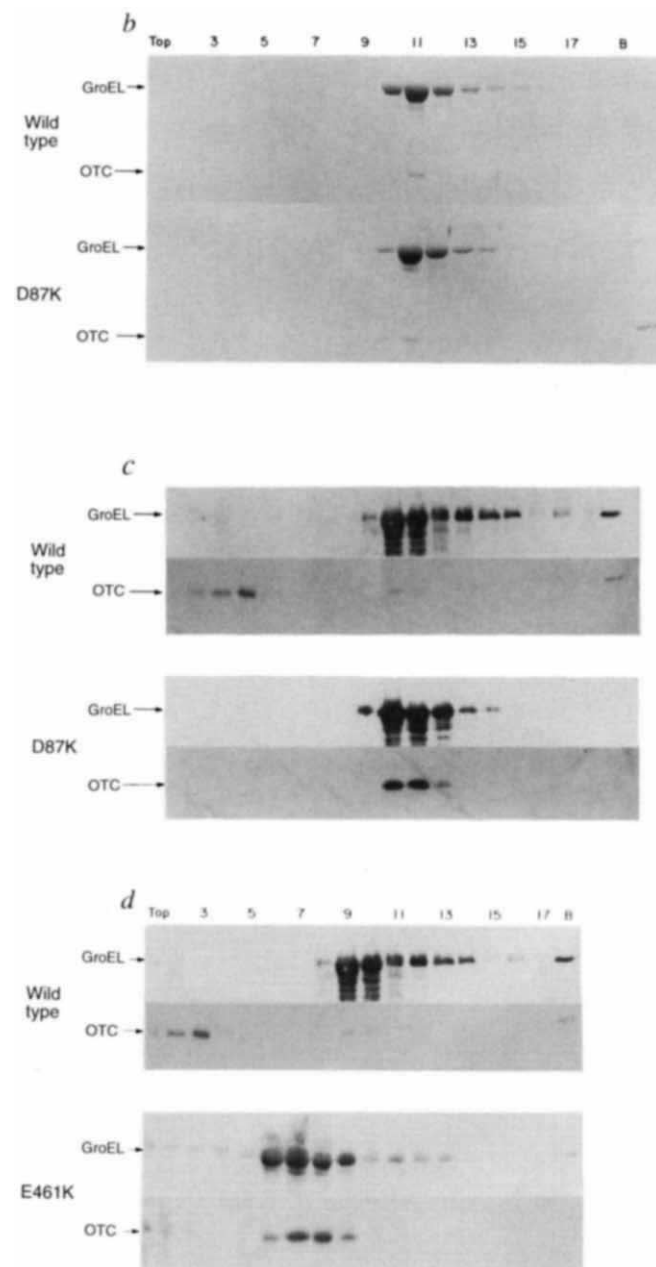


**FIG. 4** The D87K mutation, in the equatorial domain, abolishes the ATPase activity of the complex and blocks release of a substrate polypeptide. An equatorial domain mutant, E461K, also has a deficiency of polypeptide release and displays aberrant migration in a sucrose gradient. **a**, ATPase activity of wild type and D87K mutant GroEL. No ATPase activity is detected with the mutant complex, with or without GroES, even at elevated ATP concentration. **b**, Binding of unfolded OTC to wild type and D87K GroEL, monitored by sucrose gradient centrifugation. Unfolded OTC may bind somewhat less efficiently to D87K, as indicated by the recovery of some aggregated OTC from the bottom of the gradient. **c**, Release and folding of OTC from the binary complexes by GroES and Mg-ATP. OTC is efficiently released from wild type GroEL and migrates as an assembled trimer near the top of a sucrose gradient, while no release occurs from the binary complex with D87K. **d**, OTC release and folding from wild-type and E461K binary complexes by GroES and Mg-ATP. Wild type releases OTC and allows it to fold and assemble into active trimers (top panel); E461K does not (lower panel). E461K does not move so far in the gradient (~15S) but still retains the bound OTC. Despite its altered migration, E461K appears as a double ring in the electron microscope (W. Baumeister, unpublished result).

**METHODS.** **a**, ATPase activities were measured as described<sup>22</sup> using Malachite green to quantify released inorganic phosphate. Either 1 mM ATP (standard conditions; four bars on left) or 10 mM ATP (two bars on right) was added to start the reactions, and they were incubated at 37 °C for 10 min and terminated with the acidic Malachite green reagent. The effect of GroES on ATPase activity was tested by adding a twofold molar excess of the cochaperonin before adding ATP. **b**, Binding of unfolded OTC by wild-type and D87K GroEL was determined as for Fig. 2a, except that detection was by directly staining the polyacrylamide gel with Coomassie brilliant blue. Note the OTC recovered in the pellet fraction (labelled 'B') in the D87K gel (lower panel). **c**, Release of OTC from the binary complex with wild type or D87K GroEL was monitored essentially as described<sup>10</sup>. Briefly, the binary complexes were recovered from fraction 11 of the gradients in **b** by concentration and buffer exchange (to remove excess sucrose) on a Centricon concentration device (Amicon). They were incubated at 0.75 μM with 1.5 μM GroES and 5 mM ATP in reconstitution buffer (50 mM Tris-HCl, pH 7.5, 5 mM KCl, 12 mM MgCl<sub>2</sub>) for 30 min at 37 °C, diluted 10-fold with 20 mM CDTA and 50 mM Tris-HCl, pH 7.5, and separated on 10–30% sucrose gradients as in **a**. Fractions were precipitated with trichloroacetic acid and resuspended in SDS sample buffer, as were any aggregates in the bottoms of the gradient tubes, and electrophoresed as before. Gels were electroblotted and reacted sequentially with anti-OTC and anti-GroEL antisera, using chemiluminescence as before. Only the relevant portions of the films are shown. Note the absence of anti-OTC-reactive material in the OTC

On the other hand, our ability to produce ternary complexes of GroEL–OTC–GroES in the presence of ADP (Y.K. and J. Weissman, unpublished results) would argue that there is not a strict mutual exclusion.

Mutations at additional apical domain sites affected GroES binding but not polypeptide binding. These mutants were defective in polypeptide release. One such mutant, N265A, lies immediately adjacent to two sites where mutational changes lead to defects of both polypeptide and GroES binding (Table 1 and Fig. 3c). By contrast, other mutations that affect GroES binding, G337S/I349E and D261K, lie on the outer circumference of the cylinder, remote from the inside apical surface (Fig. 3c).



trimer position (fractions 2–4) of the D87K gradient. **d**, Binary complexes were prepared and OTC release assayed as in **b** and **c**, respectively. Protein blots were reacted sequentially with the antisera as before. The migration position of E461K was reproducible when fraction 7 was recovered and separated again (not shown). Likewise, E461K alone (without bound OTC) also moved to this position when treated with Mg-ATP or MgATP-γS (not shown).

ATPase activity was affected in several of the mutants (Table 1 and Fig. 4a). Two mutants mapping to a site in the equatorial domain, D87K and D87N, had no ATPase activity under any conditions (Fig. 4a). In equilibrium dialysis, wild type and D87K bound ATP- $\gamma$ S with  $K_D$  values of 0.33 and 0.81 mM, respectively (not shown). Thus, although there may be a relatively small decrease in affinity of the mutant complex for ATP- $\gamma$ S, there is a complete loss of ATP hydrolytic activity. The affected residue, present in all chaperonins in the conserved motif, GDGTT, lies at the opening to a pocket in the upper surface of the equatorial domain that has a volume sufficient to accommodate ATP (Fig. 3d). Indeed, structural analysis in progress on ATP-containing crystal forms reveals that ATP occupies this site (D. Boisvert, J. Wang, Z. Otwinowski, A.L.H. and P. Sigler, unpublished results).

Polypeptide binding by D87K was mildly reduced (Fig. 4b), but, more strikingly, there was complete block of polypeptide release (Fig. 4c). In particular, when binary GroEL(D87K)-OTC complex (fraction 11 in Fig. 4b) was concentrated and incubated with GroES and Mg-ATP, then assayed for reconstitution of OTC enzymatic activity, less than 1% was recovered, as compared with 80–90% recovery for wild-type complex (not shown). When reaction products were fractionated on a sucrose gradient, OTC subunits remained in the D87K-containing fraction at 20S (Fig. 4c). The observation of failed release contrasts with a previous observation that addition of GroES and the non-hydrolysable ATP analogue, AMP-PNP, to wild-type GroEL-OTC binary complex resulted in the release of OTC as misfolded monomer<sup>10</sup>. Indeed, this latter experiment would suggest that ATP binding may be sufficient to direct polypeptide release. Thus, a possible explanation for the behaviour of D87K is that it cannot properly transduce ATP binding into a change in affinity for polypeptide that is sufficient to bring about its release. A further implication of the AMP-PNP experiment is that hydrolysis may be required for resetting the complex to a state with high affinity for polypeptide, as has been observed for Hsp70 proteins<sup>15</sup>. For GroEL, such recycling is apparently obligatory for the multiple steps of binding and release necessary for efficient folding of a substrate peptide to its native state.

Single-residue substitutions in the intermediate domain exert global effects on chaperonin function (I150E, S151V, A152E, A383E, A405E and A406E in Table 1; and Fig. 3). All of the mutants were devoid of ATPase activity, despite their location outside the ATP-binding pocket (Fig. 3d), and all but A405E could not bind GroES. A152E could not bind polypeptide, and binding by A406E was reduced. Although the mutants at 150, 151, 383 and 405 could bind polypeptide, release was not observed upon addition of Mg-ATP and GroES. The effects of this group of mutants, involving functions associated with both apical and equatorial domains, suggest that the intermediate domain can act as a functional link mediating critical allosteric interactions between equatorial and apical domains<sup>16–19</sup> (Fig. 3a, d). Because residues in the intermediate domain make non-covalent contacts with the apical domain of the neighbouring subunit, this region may also transduce changes between subunits<sup>7</sup>.

Allosteric behaviour that crosses the equatorial interface between rings may also be required for polypeptide release, supported by results with the mutant E461K. Glu 461 resides in the underside of the equatorial domain (Fig. 3e) and reaches across the equator to form an electrostatic contact with Arg 452 in the dyad-related subunit (Fig. 4b in the accompanying paper<sup>7</sup>). Surprisingly, mutation at this equatorial site has no effect on ATPase activity but affects GroES binding and blocks polypeptide release (Table 1 and Fig. 4d), events involving the apical domain.

In summary, our structural<sup>7</sup> and functional studies suggest that non-native polypeptides bind inside the central channel of GroEL by hydrophobic interaction with residues in the inner face of the apical domains. Both GroES binding to the apical

domains of one or both rings<sup>20,21</sup>, and ATP binding in the underlying equatorial domain contribute to allosteric conformational changes that enable productive release of non-native polypeptide. As shown recently, the released polypeptide partitions itself between folding into the native form or binding to another GroEL molecule<sup>4,6</sup>. This cycle of release and kinetic partitioning continues until folding is complete. □

Received 30 August; accepted 15 September 1994.

1. Gething, M. J. & Sambrook, J. *Nature* **355**, 33–45 (1992).
2. Hendrick, J. P. & Hartl, F.-U. *A. Rev. Biochem.* **62**, 349–384 (1993).
3. Horwich, A. L. & Willison, K. *Phil. Trans. R. Soc. Lond.* **B339**, 313–326 (1993).
4. Todd, M. J., Viitanen, P. V. & Lorimer, G. H. *Science* **265**, 659–666 (1994).
5. Gao, Y. et al. *J. Cell Biol.* **125**, 989–996 (1994).
6. Weissman, J. S., Kashi, Y., Fenton, W. A. & Horwich, A. L. *Cell* **78**, 693–702 (1994).
7. Braig, K. et al. *Nature* **371**, 578–586 (1994).
8. Horwich, A. L., Low, K. B., Hirshfield, I. N. & Furtak, K. *Cell* **74**, 909–917 (1993).
9. Cheng, M. Y. et al. *Nature* **337**, 620–625 (1989).
10. Zheng, X., Rosenberg, L. E., Kalousek, F. & Fenton, W. A. *J. Biol. Chem.* **268**, 7489–7493 (1993).
11. Martin, J. et al. *Nature* **352**, 36–42 (1991).
12. Landry, S. et al. *Nature* **355**, 455–457 (1992).
13. Chen, S. et al. *Nature* **371**, 261–264 (1994).
14. Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. & Gierasch, L. M. *Nature* **364**, 255–258 (1993).
15. Palleros, D. R., Reid, K. L., Shi, L., Welch, W. J. & Fink, A. L. *Nature* **365**, 664–666 (1993).
16. Gray, T. E. & Fersht, A. R. *FEBS Lett.* **292**, 254–258 (1991).
17. Bochkareva, E. S., Lissin, N. M., Flynn, G. C., Rothman, J. E. & Girshovich, A. S. *J. Biol. Chem.* **267**, 6796–6800 (1992).
18. Jackson, G. S. et al. *Biochemistry* **32**, 2554–2563 (1993).
19. Langer, T., Pfeifer, G., Martin, J., Baumeister, W. & Hartl, F. U. *EMBO J.* **11**, 4757–4765 (1992).
20. Azem, A., Kessel, M. & Goloubinoff, P. *Science* **265**, 653–656 (1994).
21. Schmidt, M. et al. *Science* **265**, 656–659 (1994).
22. Trent, J. D., Nimmesgern, E., Wall, J. S., Hartl, F. U. & Horwich, A. L. *Nature* **354**, 490–493 (1991).

ACKNOWLEDGEMENTS. We thank A. Hack for technical assistance and D. Boisvert, K. Braig, A. Joachimiak, Z. Otwinowski, P. Sigler and J. Weissman for discussion. This work was supported by grants from the NIH. Y.K. was a Fogarty Fellow.

## Ribozyme-mediated repair of defective mRNA by targeted *trans*-splicing

Bruce A. Sullenger\* & Thomas R. Cech†

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA

RIBOZYMES can be targeted to cleave specific RNAs<sup>1–8</sup>, which has led to much interest in their potential as gene inhibitors<sup>3,9,10</sup>. Such *trans*-cleaving ribozymes join a growing list of agents that stop the flow of genetic information<sup>11,12</sup>. Here we describe a different application of ribozymes for which they may be uniquely suited. By targeted *trans*-splicing, a ribozyme can replace a defective portion of RNA with a functional sequence. The self-splicing intron from *Tetrahymena thermophila*<sup>13</sup> was previously shown to mediate *trans*-splicing of oligonucleotides *in vitro*<sup>14,15</sup>. As a model system for messenger RNA repair, this group I intron was re-engineered to regenerate the proper coding capacity of short, truncated *lacZ* transcripts. *Trans*-splicing was efficient *in vitro* and proceeded in *Escherichia coli* to generate translatable *lacZ* messages. Targeted *trans*-splicing represents a general means of altering the sequence of specified transcripts and may provide a new approach to the treatment of many genetic diseases.

The *Tetrahymena* self-splicing group I intron efficiently splices itself from *lacZ* transcripts in *E. coli*, generating mRNA for the translation of the  $\alpha$ -complement of  $\beta$ -galactosidase ( $\beta$ -gal) (Fig. 1a)<sup>16,17</sup>. We tested whether the ribozyme could perform a similar reaction *in trans*. This system consists of two RNA molecules

\* Present address: Departments of Experimental Surgery and Genetics, Duke University Medical Center, Durham, North Carolina 27710, USA.

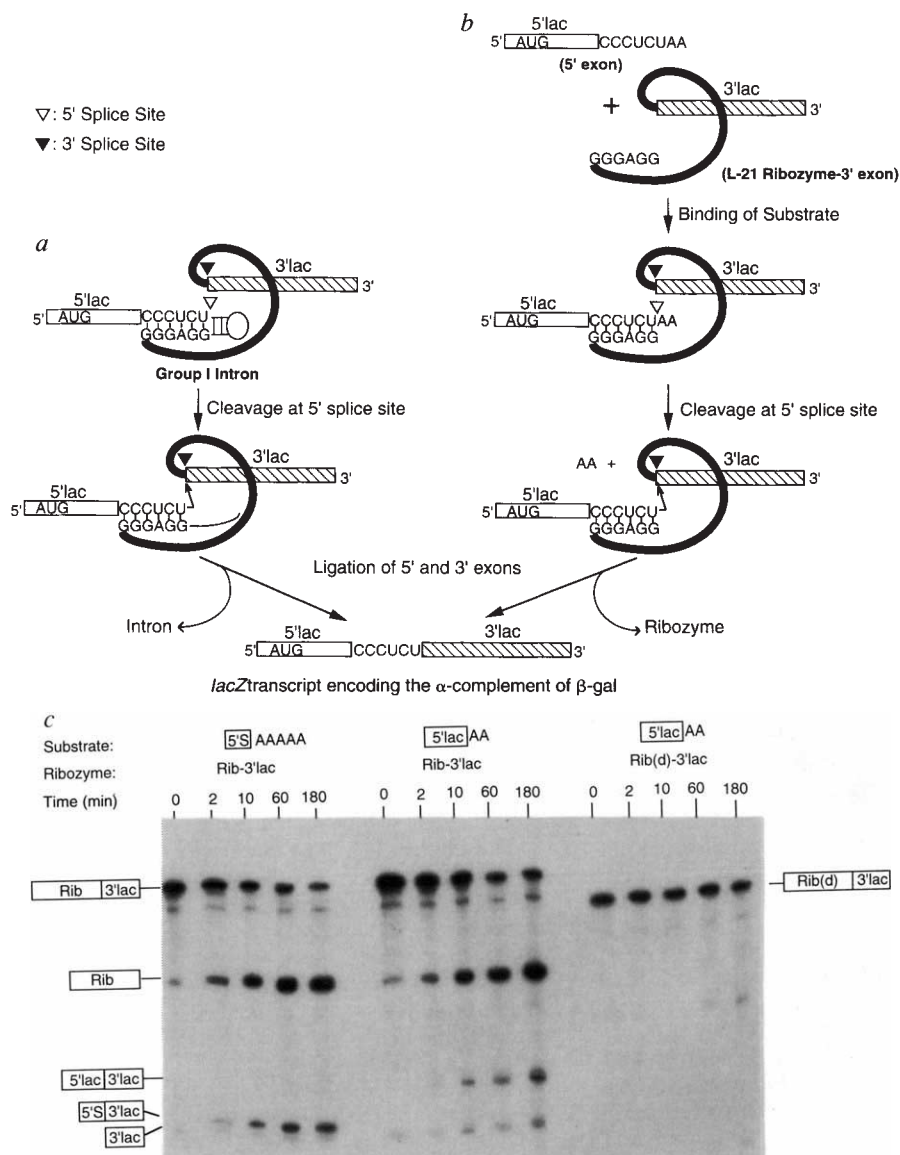
† To whom correspondence should be addressed.



FIG. 1 *Cis*- and *trans*-splicing of *lacZ* transcripts. *a*, *Cis*-splicing experiments of ref. 16, in which the self-splicing intron from *Tetrahymena* and flanking ribosomal RNA exon sequences were inserted within the *lacZ* gene. Self-splicing is initiated by attack of an intron-bound guanosine at the 5' splice site (▽). Ligation of the freed 5' exon to the 3' splice site (▼) then yields mRNA with an open reading frame encoding the  $\alpha$ -complement of  $\beta$ -gal. Translation of these RNAs results in the production of a protein with a few extra amino acids near its amino terminus. The  $\alpha$ -complement of  $\beta$ -gal can tolerate such extra amino acids without destroying its activity<sup>20</sup>.

*b*, Proposed *trans*-splicing to correct truncated *lacZ* transcripts. The 5' exon RNA (5'*lac*AA) consists of the first 31 nucleotides of the *lacZ* transcript, the recognition sequence (CCCUCU) which is complementary to the 5' exon-binding site of the ribozyme (GGGAGG), and two adenosines. The sequence of this 39 nucleotide substrate is: 5'-AGG AAA CAG CUA UGA CCA UGA UUA CGG AUU CCC CUC UAA-3' (where the AUG initiation codon is underlined). The 393-nucleotide ribozyme is followed by a 3' exon which contains 23 nucleotides derived from the 3' exon of the pre-rRNA of *Tetrahymena* and 200 nucleotides of *lacZ*. The L-21 form of the ribozyme used lacks the first 21 nucleotides of the 5' end of the intron from which it is derived<sup>18</sup>.

*Trans*-splicing results in the production of corrected *lacZ* transcripts which contain an open reading frame for the translation of the  $\alpha$ -complement of  $\beta$ -gal. Once again, a few extra amino acids will be contained in the protein which should not affect its activity. *c*, *Trans*-splicing *in vitro*. To generate body-labelled ribozyme-3' exon chimaeric RNAs (Rib-3'*lac* or Rib(d)-3'*lac*), the plasmid pT7L-21 (ref. 18) or pT7L-21del was digested with *Nde*I, and T7 RNA polymerase was used to synthesize run-off transcripts in the presence of [ $\alpha$ -<sup>32</sup>P]ATP. The inactive Rib(d) ribozyme was generated by deleting nucleotides 237-330 of the intron, which include the G-binding site<sup>23</sup>. The labelled ribozyme-3' exon RNAs (200 nM) were preheated in reaction buffer (50 mM HEPES (pH 7.0), 150 mM NaCl, and 5 mM MgCl<sub>2</sub>) at 50 °C for 5 min and then equilibrated at 37 °C for 2 min. The 13 nucleotide (5'SAAAAA: GGCCUCUAAAAA) or 39 nucleotide (5'*lac*AA: b) substrate RNAs (1  $\mu$ M) and GTP (100  $\mu$ M) were preheated to 37 °C and added to the ribozymes to start the reactions, which proceeded at 37 °C. Portions containing one fifteenth of the



(Fig. 1b): (1) 5'*lac* RNA, which contains a ribosome binding site, the first 21 nucleotides of the *lacZ* coding sequence and the 5' splice site sequence CCCUCU/AA (where the solidus represents the site of cleavage), and (2) ribozyme-3'*lac* RNA, comprised of the L-21 form of the *Tetrahymena* intron<sup>18</sup> covalently linked to a 3' exon that encodes 67 amino acids of the  $\alpha$ -complementation fragment of  $\beta$ -gal. For *trans*-splicing to correct these truncated *lacZ* transcripts, the ribozyme must recognize the defective 5'*lac* RNA by base-pairing, cleave off the two additional adenosine residues, hold onto the 5' exon cleavage product by base pairing to the 5' exon-binding site of the ribozyme, and ligate the *lacZ* 3' exon sequence onto the cleaved 5' product to give the correct translational reading frame (Fig. 1b).

*In vitro*, the ribozyme can quickly and accurately *trans*-splice its *lacZ* 3' exon onto the truncated *lacZ* 5' exon (Fig. 1c). Radio-labelled ribozyme-*lacZ* 3' exon RNA (Rib-3'*lac*) was allowed to react with an excess of unlabelled 5' exon substrate RNA. Rib-3'*lac* RNA was converted to free L-21 ribozyme (Rib) plus ligated exons (5'S-3'*lac* or 5'*lac*-3'*lac*). The reaction with the 39 nucleotide *lacZ* 5' exon substrate, called 5'*lac*AA, occurred with

reactions were removed at the times indicated and added to an equal volume of 10 mM EDTA to stop the reactions. Reaction products were analysed on a 4% polyacrylamide gel containing 8 M urea.

a half-time of 13 min, and 85-90% of the precursor was converted to *trans*-spliced RNA plus free ribozyme. A shorter 13 nucleotide 5' exon substrate, called 5'SAAAAA, reacted with similar rate and efficiency. An inactive version of the L-21 ribozyme (Rib(d)), generated by deleting part of its catalytic core, was not able to perform *trans*-splicing (Fig. 1c).

Direct RNA sequencing of the *trans*-spliced product (5'*lac*-3'*lac* in Fig. 1c) demonstrated that the ribozyme correctly removed the ultimate and penultimate 3' adenosine nucleotides from the substrate RNA 5'*lac*AA, and accurately spliced the *lacZ* 3' exon sequences onto the cleaved 5' exon product (data not shown). The repaired *lacZ* transcript contained the proper open reading frame for expression of the  $\alpha$ -complement of  $\beta$ -gal.

To determine if *trans*-splicing could correct truncated *lacZ* transcripts in *E. coli*, we constructed the plasmid pT/S (Fig. 2a). Both the ribozyme-*lacZ* 3' exon RNA and the truncated *lacZ* 5' exon RNA, described in the *in vitro* experiments above, can be expressed from pT/S in *E. coli* containing T7 RNA polymerase. The plasmid pT/Sd (Fig. 2a) is identical except that it contains the inactive derivative of the ribozyme.

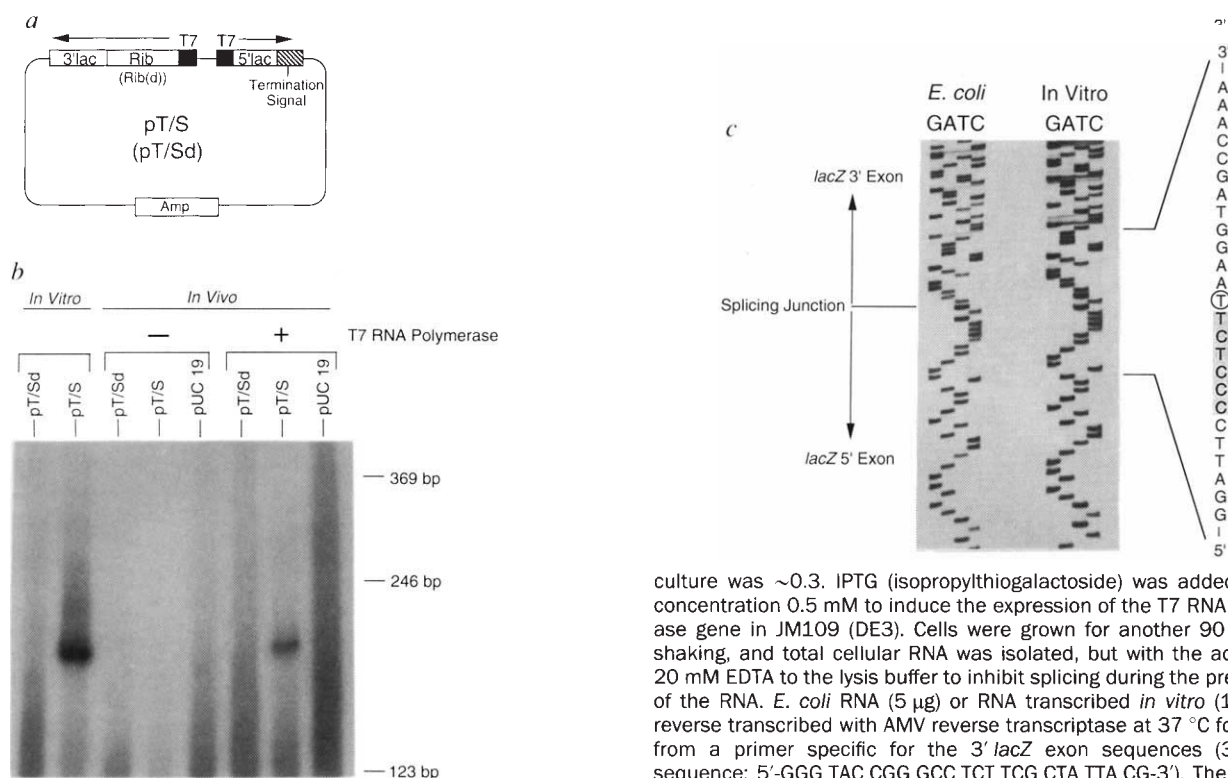


FIG. 2 Targeted *trans*-splicing in *E. coli*. *a*, *Trans*-splicing plasmids. pT/S contains two T7 RNA polymerase promoters. One directs the expression of the 39 nucleotide 5' exon sequence of *lacZ*, and the other directs the expression of the L-21 ribozyme-*lacZ* 3' exon chimaeric RNA; both are as shown in Fig. 1*b*, except the 5'*lac* sequence is followed by nine adenosines and a transcription termination signal for T7 RNA polymerase. pT/Sd is identical to pT/S except that it encodes the inactive ribozyme Rib(d). pT/S and pT/Sd were created by cloning oligonucleotides containing the T7 RNA polymerase promoter, the 5' exon *lacZ* sequence, and the T7 transcription termination signal into pT7L-21 and pT7L-21del. *b*, RT-PCR analysis of *trans*-spliced RNA products generated *in vitro* or in *E. coli*. The pT/S and pT/Sd plasmids were *in vitro* transcribed with T7 RNA polymerase for 1 h at 37 °C. In addition, these plasmids were electroporated into the *E. coli* strains JM109 (ref. 20) and JM109 (DE3) (Promega, Madison, WI); the latter strain contains a chromosomal copy of the T7 RNA polymerase gene transcribed by the *lac* UV5 promoter. pUC19 was also electroporated into these *E. coli* strains to serve as a control; it produces *lacZ* mRNA but with a different sequence, which should not be amplified by RT-PCR (see below). Bulk transformed cells were grown in LB media at 37 °C until the A<sub>600</sub> of the

The plasmids pT/S, pT/Sd and pUC19 were separately introduced into *E. coli* strains that produce or do not produce T7 RNA polymerase. RNAs were isolated from cells stably transduced with these plasmids. In addition, the pT/S and pT/Sd plasmids were transcribed *in vitro* with T7 RNA polymerase, and the resulting RNAs were also analysed. To determine if *trans*-splicing had occurred in any of the RNA samples, reverse transcription and polymerase chain reaction (RT-PCR) analyses were done using one primer specific for the truncated *lacZ* 5' exon and the other primer specific for the 3' exon (Fig. 2*b*). An amplified DNA fragment of the expected size (200 base pairs) was generated from the RNA isolated from cells that contained both pT/S and T7 RNA polymerase, and from the RNA generated by *in vitro* transcription of pT/S. No such RT-PCR product was generated from RNA from cells that contained pT/Sd, which encodes the defective ribozyme, or from RNA from cells that contained the active pT/S plasmid but did not express T7 RNA polymerase (Fig. 2*b*). Therefore, *trans*-splicing in *E. coli* was inferred to depend on expression of the active ribozyme.

To confirm that the amplified DNA fragments were generated from *lacZ* transcripts, the fragments were cloned and sequenced

culture was ~0.3. IPTG (isopropylthiogalactoside) was added to final concentration 0.5 mM to induce the expression of the T7 RNA polymerase gene in JM109 (DE3). Cells were grown for another 90 min with shaking, and total cellular RNA was isolated, but with the addition of 20 mM EDTA to the lysis buffer to inhibit splicing during the preparation of the RNA. *E. coli* RNA (5 µg) or RNA transcribed *in vitro* (1 µg) was reverse transcribed with AMV reverse transcriptase at 37 °C for 30 min from a primer specific for the 3'*lacZ* exon sequences (3' primer sequence: 5'-GGG TAC CGG GCC TCT TCG CTA TTA CG-3'). The samples were then treated with RNase A for 30 min at 37 °C. The resulting cDNAs were PCR amplified for 30 cycles using this 3' primer as well as a 5' primer (5'-CGG GAT CCA TGA TTA CGG ATT CCC CT-3'), which is specific for the 5'*lacZ* exon present on pT/S because it ends with CCCT. Thus, the wild-type *lacZ* transcript is not amplified. [ $\alpha$ -<sup>32</sup>P]dCTP was added to the PCR reactions to label the amplified products which were separated on a 6% polyacrylamide gel. If *trans*-splicing occurred correctly, an amplified DNA fragment of 200 base pairs would be expected. A 123 base pair DNA ladder visualized by staining provided molecular mass markers. *c*, Sequence of *trans*-splicing PCR products. The PCR amplified products shown in *b* (pT/S *in vitro* and in *E. coli* expressing T7 RNA polymerase) were gel isolated and cloned into pUC19. The inserts of three different plasmids from *in vitro* samples and three from *E. coli* samples were sequenced by the dideoxynucleotide method. All sequences were identical, and representative ones from *E. coli* and *in vitro* transcription derived samples are shown. Sequence ladders were separated on an 8% polyacrylamide gel containing 8 M urea. The expected sequence around the splicing junction is also shown, with the ribozyme recognition sequence shaded and the first nucleotide of the 3' exon circled.

(Fig. 2*c*). The amplified products encode the corrected *lacZ* transcript, demonstrating that the ribozyme correctly joined the 5' and 3' *lacZ* exons both *in vitro* and in *E. coli*.

Experiments to test translation of the repaired *lacZ* transcripts *in vivo* used *E. coli* that have no endogenous  $\beta$ -gal activity. Cells containing pT/S, pT/Sd or pUC19 were transformed with a second, compatible plasmid that expresses T7 RNA polymerase when induced by heat shock<sup>19</sup>. Cells containing pUC19 had high  $\beta$ -gal activity regardless of the presence of T7 RNA polymerase (Table 1). This was expected because in pUC19 the expression of the *lacZ* gene is under the control of the *lac*UV5 promoter<sup>20</sup>, and thus is not dependent on T7 RNA polymerase. In contrast, *E. coli* containing pT/S showed  $\beta$ -gal activity only in the presence of T7 RNA polymerase. The level of activity was 0.5–1.0% of that measured in cells containing pUC19 (Table 1). Cells containing pT/Sd, which produces the defective ribozyme, did not accumulate significant levels of  $\beta$ -gal in the presence or absence of T7 RNA polymerase. Therefore, the repaired *lacZ* transcripts are apparently translated in *E. coli*, even though transcription and translation are normally coupled in bacteria.



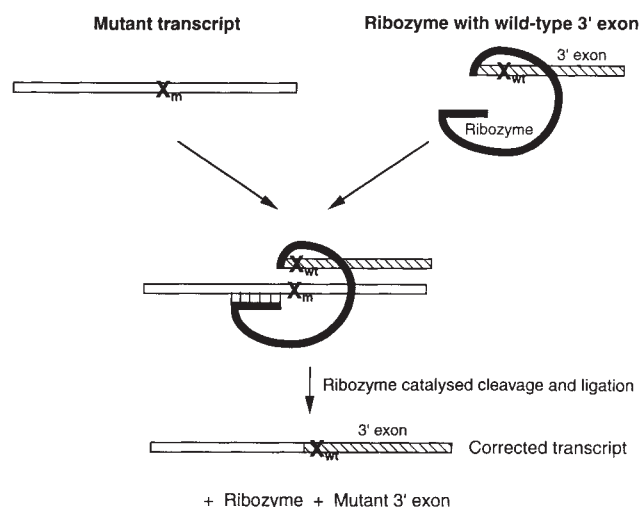


FIG. 3 Scheme for correction of mutant transcripts by targeted trans-splicing.  $X_m$ , Point mutation at nucleotide X;  $X_{wt}$ , wild-type nucleotide.

RNA isolated from these cells was also analysed for trans-splicing products using the RT-PCR method described above. Amplified fragments of the correct size were generated with RNA samples from the cells containing pT/S only when the plasmid encoding T7 RNA polymerase was also present (data not shown). In contrast, no amplified fragment was generated with RNA samples from cells containing either pT/Sd or pUC19. As before, the lack of signal from cells containing pUC19 is due to the primer DNA sequence, which made the RT-PCR specific for the trans-spliced *lacZ* mRNA. This RNA analysis, coupled with the observation that  $\beta$ -gal activity was generated in *E. coli*, provides a particularly convincing demonstration that the ribozyme functions *in vivo*.

Targeted trans-splicing can potentially be adapted to correct a broad array of mutant transcripts. Group I introns recognize their splice sites by complementary base pairing to their 5' exon-binding site<sup>15,21</sup>. The only specific sequence requirement is a single uridine preceding the cleavage site, so a trans-splicing group I intron can be designed to cleave a targeted transcript at any available uridine residue upstream of a mutation. There is no specific sequence requirement for a 3' exon in such reactions, although sequences capable of forming the P10 pairing with the ribozyme may be advantageous<sup>13</sup>.

The relatively short 5' exon-binding site of group I introns (<9 nucleotides among known examples) would appear to limit the sequence specificity of the ribozyme for its target. However, colocalization of a ribozyme with its target may provide the requisite specificity<sup>22</sup>. Furthermore, much is known about the energetics of substrate binding to these ribozymes<sup>23</sup>, so several logical approaches exist for redesigning ribozymes to increase their sequence specificity<sup>24</sup>.

TABLE 1 Trans-spliced RNAs are translated *in vivo*

| Plasmid | T7 RNA polymerase* | $\beta$ -gal activity (units) <sup>†</sup> |
|---------|--------------------|--------------------------------------------|
| pUC19   | —                  | 20,100 $\pm$ 3200                          |
|         | +                  | 16,400 $\pm$ 500                           |
| pT/S    | —                  | 0                                          |
|         | +                  | 107 $\pm$ 2                                |
| pT/Sd   | —                  | 0                                          |
|         | +                  | 4 $\pm$ 2                                  |

Because JM109 (DE3) cells have endogenous  $\beta$ -gal activity (data not shown), experiments to test the translation of trans-spliced *lacZ* mRNA were done in JM109. JM109 cells containing pUC19, pT/S, or pT/Sd were transduced by electroporation with the pACYC-derived plasmid pGP1-2 (ref. 19), which encodes T7 RNA polymerase. The pT/S plasmids are pUC derivatives and contain a Col E1 origin of replication, which is compatible with the P15A origin of pGP1-2 (ref. 19). Bulk transfected and mock transfected cells (10 ml) were grown at 30 °C until the absorbance at 590 nm of the culture was about 0.4, heat shocked at 42 °C for 30 min to induce the T7 RNA polymerase, and then grown for 90 min at 37 °C with shaking<sup>29</sup>. The *E. coli* were then pelleted and assayed for  $\beta$ -gal activity using a colorimetric assay in which o-nitrophenyl- $\beta$ -D-galactoside (ONPG) is cleaved to produce o-nitrophenol<sup>30</sup>. A control sample from JM109 cells transfected with pBR322 was used to zero the spectrophotometer.

\* +, Cells transfected with pGP1-2; —, mock transfected.

<sup>†</sup> Mean  $\pm$  s.d. of triplicate samples. Unit = 1,000  $\times$  A<sub>420</sub> ml<sup>-1</sup> h<sup>-1</sup>.

In this era of gene mapping and human genome sequencing, the molecular basis for an increasing number of inherited diseases is being discovered. Conventional gene therapy attempts to correct a genetic deficiency by transferring a wild-type complementary DNA version of a gene, usually under the control of a heterologous promoter, to cells harbouring a defective copy of the gene. One major obstacle for implementing such treatments is our inability to recapitulate faithfully the normal expression pattern of endogenous genes after gene transfer<sup>25</sup>. Targeted trans-splicing can potentially solve this problem. Endogenous transcripts that are truncated or contain point mutations can be repaired (Fig. 3), while the natural regulated expression of the genes is maintained.

Trans-splicing ribozymes may also act as effective antiviral agents by changing the sequence of viral RNAs so that they encode dominant negative versions of viral proteins<sup>26,27</sup>.

Although there are many challenges in development of any gene therapy protocol, the results presented here suggest that ribozymes may have therapeutic utility in a direction besides those envisioned previously. Future goals include increasing the efficiency of trans-splicing, applying it to full-length endogenous transcripts, and testing it in mammalian cells. Many group I introns function in eukaryotic nuclei in nature, and a group I intron introduced into an RNA polymerase II transcript has recently been shown to function in yeast<sup>28</sup>. Thus, there is potential for group I ribozymes to promote trans-splicing in mammalian cells. □

Received 12 July; accepted 24 August 1994.

1. Zaug, A. J., Been, M. D. & Cech, T. R. *Nature* **324**, 429–433 (1986).
2. Uhlenbeck, O. C. *Nature* **328**, 596–600 (1987).
3. Haseloff, J. & Gerlach, W. L. *Nature* **334**, 585–591 (1988).
4. Feldstein, P. A., Buzayan, J. M. & Bruening, G. *Gene* **2**, 53–61 (1989).
5. Hampel, A., Tritz, R., Hicks, M. & Cruz, P. *Nucleic Acids Res.* **18**, 299–304 (1990).
6. Chowrira, B. M. & Burke, J. M. *Biochemistry* **30**, 8518–8522 (1991).
7. Forster, A. C. & Altman, S. *Science* **249**, 783–786 (1990).
8. Perrotta, A. T. & Been, M. D. *Biochemistry* **31**, 16–21 (1992).
9. Cech, T. R. *J. Am. med. Assoc.* **260**, 3030–3034 (1988).
10. Rossi, J. J. *Curr. Opin. Biotech.* **3**, 3–7 (1992).
11. Gilboa, E. & Smith, C. *Trends Genet.* **10**, 139–144 (1994).
12. Yu, M., Poeschla, E. & Wong-Staal, F. *Gene Ther.* **1**, 13–26 (1994).
13. Cech, T. R. *Rev. Biochem.* **59**, 543–568 (1990).
14. Inoue, T., Sullivan, F. X. & Cech, T. R. *Cell* **43**, 431–437 (1985).
15. Been, M. D. & Cech, T. R. *Cell* **47**, 207–216 (1986).
16. Price, J. V. & Cech, T. R. *Science* **228**, 719–722 (1985).
17. Waring, R. B. et al. *Cell* **40**, 371–380 (1985).
18. Zaug, A. J., Grosshans, C. A. & Cech, T. R. *Biochemistry* **27**, 8924–8931 (1988).

19. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1074–1078 (1985).
20. Yanisch-Perron, C., Vieira, J. & Messing, J. *Gene* **33**, 103–119 (1985).
21. Waring, R. B., Townner, P., Minter, S. & Davies, R. W. *Nature* **321**, 133–139 (1986).
22. Sullenger, B. A. & Cech, T. R. *Science* **262**, 1566–1569 (1993).
23. Cech, T. R., Herschlag, D., Piccirilli, J. A. & Pyle, A. M. *J. biol. Chem.* **267**, 17479–17482 (1992).
24. Herschlag, D. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6921–6925 (1991).
25. Morgan, R. A. & Anderson, W. F. *Rev. Biochem.* **62**, 191–217 (1993).
26. Malim, M. H., Bohnlein, E., Hauber, J. & Cullen, B. R. *Cell* **58**, 205–214 (1989).
27. Tronam, D., Feinberg, M. B. & Baltimore, D. *Cell* **59**, 113–120 (1989).
28. Ford, E. & Ares, M. Jr *Proc. natn. Acad. Sci. U.S.A.* **91**, 3117–3121 (1994).
29. Tabor, S. in *Current Protocols in Molecular Biology* (eds Ausubel, F. A. et al.) 16.2.1–16.2.11 (Greene/Wiley, New York, 1990).
30. Miller, J. in *Experiments in Molecular Genetics* 352–355 (Cold Spring Harbor Laboratory Press, New York, 1972).

ACKNOWLEDGEMENTS. We thank T. McConnell, J. Doudna, A. Zaug, F. Murphy, T. Campbell, S. Strobel and R. Rempel for useful discussions and C. Grosshans and A. Gooding for technical support. B.A.S. was a Damon Runyon-Walter Winchell Cancer Research Fund Postdoctoral Fellow, and is supported by a grant from the NIH; T.R.C. is an Investigator of the Howard Hughes Medical Institute and an American Cancer Society Professor.

# Crystal structure of an RNA bacteriophage coat protein–operator complex

Karin Valegård\*, James B. Murray†, Peter G. Stockley†, Nicola J. Stonehouse† & Lars Liljas\*

\* Department of Molecular Biology, Uppsala University, Box 590, S-751 24 Uppsala, Sweden

† Department of Genetics, University of Leeds, Leeds LS2 9JT, UK

THE RNA bacteriophage MS2 is a convenient model system for the study of protein–RNA interactions. The MS2 coat protein achieves control of two distinct processes—sequence-specific RNA encapsidation and repression of replicase translation—by binding to an RNA stem–loop structure of 19 nucleotides containing the initiation codon of the replicase gene. The binding of a coat protein dimer to this hairpin shuts off synthesis of the viral replicase<sup>1</sup>, switching the viral replication cycle to virion assembly rather than continued replication. The operator fragment alone can trigger self-assembly of the phage capsid at low protein concentrations and a complex of about 90 RNA operator fragments per protein capsid has been described<sup>2</sup>. We report here the crystal structure at 3.0 Å resolution of a complex between recombinant MS2 capsids and the 19-nucleotide RNA fragment. It is the first example of a structure at this resolution for a sequence-specific protein–RNA complex apart from the transfer RNA synthetase complexes<sup>3–5</sup>. The structure shows sequence-specific interactions between conserved residues on the protein and RNA bases essential for binding.

The protein shell of MS2 is formed from 180 identical protein subunits, each comprising 129 amino acids, arranged in an icosahedral  $T=3$  surface lattice. The icosahedral asymmetric unit contains three subunits—A, B and C. The protein shell is built up of 90 dimers which have two distinct but similar conformations<sup>6</sup>: CC' dimers at the 2-fold axis and AB' dimers at the quasi-2-fold axis (Fig. 1).

A synthetic RNA fragment was constructed (see Fig. 2a for sequence) and used to prepare an MS2 coat protein–RNA complex. Only four of the 19 nucleotides are essential for coat protein recognition, provided that the Watson–Crick base pairing of the stem is maintained. The four<sup>7</sup> are the As at positions –4, –7 and –10 and the pyrimidine at position –5. In the synthetic fragment, the wild-type U was replaced by C at position –5, producing a variant with higher affinity for the coat protein<sup>8</sup>. A short description of the structure determination of the complex is given in Table 1.

The quality of the electron density map for the nucleic acid is shown in Fig. 2b. The electron density of the nucleic acid is comparable in height to that for the protein. As the RNA fragment is asymmetric, its complete conformation can be analysed only because it binds asymmetrically to the AB' dimer at the particle quasi-2-fold axis. The RNA hairpin is composed of a base-paired stem, of an A-type, containing a single bulged nucleotide and a four-nucleotide loop (Fig. 2c). The three first and the three last nucleotides of the 19-mer could not be modelled in the electron density map. However, a total of 13 nucleotides, U–12 to A +1, are clearly identified. The overall RNA hairpin structure has the shape of a crescent, where the contacts to the coat protein dimer are mainly at the narrow ends. The quasi-2-fold axis passes through the centre of the crescent, consequently the bulged A –10 and A –4 in the loop are quasi-2-fold related. The RNA segment binds across the  $\beta$ -sheet, as can be clearly seen in Fig. 3a. Electron density is also found at the crystallographic 2-fold axis, in contact with the CC' dimer. Because of the 2-fold symmetry at this position, the electron

TABLE 1 Summary of crystallographic analysis

| Resolution (Å)        | Data coverage (%) | $R_{\text{merge}}$ * |
|-----------------------|-------------------|----------------------|
| 8.9                   | 59.9              | 0.100                |
| 6.5                   | 59.5              | 0.094                |
| 5.4                   | 58.8              | 0.105                |
| 4.7                   | 58.0              | 0.110                |
| 4.2                   | 57.5              | 0.109                |
| 3.8                   | 56.6              | 0.132                |
| 3.6                   | 51.0              | 0.144                |
| 3.4                   | 33.2              | 0.166                |
| 3.2                   | 21.9              | 0.192                |
| 3.0                   | 6.7               | 0.236                |
| Total                 | 41.5              | 0.115                |
| Refinement            |                   |                      |
| Resolution (Å)        | 10.0–3.0          |                      |
| $R$ -factor†          | 0.192             |                      |
| Free $R$ -factor      | 0.209             |                      |
| Reflections           | 73,265            |                      |
| Total number of atoms | 4,433             |                      |

Recombinant MS2 coat protein was expressed in *E. coli* and capsids were purified as previously described<sup>21</sup>. Crystals of recombinant capsids were grown from hanging drops at 37 °C. The drops contained equal volumes of a stock solution and a solution containing 10 mg ml<sup>–1</sup> capsid in 0.1 M sodium phosphate pH 7.4, 0.02% (w/v) sodium azide. The stock solution contained 2.5% (w/v) polyethyleneglycol 8000 in 0.1 M sodium phosphate, pH 7.4, and 0.02% (w/v) sodium azide. The droplets were equilibrated against a solution containing 0.4 M sodium phosphate, pH 7.4, and 0.02% (w/v) sodium azide. The crystals grew in about 2 weeks to dimensions up to 1.0 mm. The space group is R32 with cell dimensions  $a = b = 288.0$  Å,  $c = 653.0$  Å,  $\alpha = \beta = 90^\circ$  and  $\gamma = 120^\circ$ . The crystals diffract to beyond 2.7 Å resolution. Data to 2.7 Å have been collected from these crystals and the structure of the empty capsid, which is very similar to the native capsid structure, has been determined and refined (K.V. and L.L., manuscript in preparation). The 19-mer RNA fragment, with the sequence 5'-ACA UGA GGA UCA CCC AUG U-3', was synthesized chemically using standard phosphoramidite chemistry<sup>22</sup> on an Applied Biosystems PCR mate 391 DNA synthesizer. The synthesis, purification and characterization procedure used was as described<sup>23</sup>. The purified RNA was desalted and concentrated with an Amicon centricon-3 to give RNA at 14 mg ml<sup>–1</sup> in 0.01% (v/v) acetonitrile and 0.24 mM ammonium acetate, pH 6.5. Crystals of the recombinant capsids were washed several times in a solution containing 0.4 M sodium phosphate pH 7.4, 5% (w/v) polyethyleneglycol 8000, RNase inhibitor (10 ml in 1 ml) and 10<sup>–4</sup> M dithiothreitol. The RNA was added to the crystal droplet in tenfold excess with a soaking time of one week. Data were collected at 5 °C on R-Axis II image systems using two different sources. Data from one crystal were collected at station 9.6 of the SERC Daresbury Synchrotron Radiation Source using wavelength 0.9 Å and data from two other crystals were obtained using a rotating anode generator at wavelength 1.54 Å. Each image covered an oscillation range of 0.50°. The images were autoindexed and processed using the program Denzo by Z. Otwinowski and scaled using the CCP4 program package. The structure was solved using molecular replacement. The refined capsid structure of bacteriophage MS2<sup>19</sup> was used as a model giving an initial  $R$ -factor of 0.290 for all reflections above  $2\sigma$  at 3.0 Å resolution. The phases were refined using the non-crystallographic symmetry. A map was computed using the CCP4 package and averaged in real space over the tenfold non-crystallographic redundancy using the program Average<sup>24</sup>. The mask was constructed using the programs MAMA (by G. Kleywegt) and O (ref. 25). A model was built in the final 3.0 Å resolution electron density map using the program O. The model, the AB' dimer–RNA complex, 13 nucleotides per dimer, and the CC' dimer–RNA complex, nine nucleotides with half occupancy, was then refined using the program XPLOR<sup>26</sup> reducing the  $R$ -factor to 0.263. Individual isotropic temperature factors were refined followed by cycles of positional refinement, resulting in an  $R$ -factor for the model of 0.216. A total of 123 ordered solvent molecules were added with further cycles of refinements. The final  $R$ -factor of the model was 0.192 against data  $>2\sigma$  from 10 Å to 3.0 Å. A small portion of the data (10%) was used to calculate the free  $R$ -factor and was never included in the refinement<sup>27</sup>.

\*  $(\sum |I - \langle I \rangle| / \sum I)$ .

†  $R$ -factor  $(\sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}})$ .



density will overlap. The asymmetry of the RNA hairpin, however, allows nine nucleotides, G' -9 to C' -1, to be built into the density close to the CC' dimer. The conformation of these nucleotides is the same as in the AB' dimer.

The structure of the coat protein subunits in complex with the 19-mer RNA is little changed from their structure in the native

virus or in the empty capsid (K.V. and L.L., manuscript in preparation). When bound to RNA, a few residues (Arg B49, Lys B57 and Asn A87) which interact directly, have slightly different conformations. Arginine B49 interacts with the two phosphate oxygens of the guanine nucleotide at -8. The side chain of lysine B57 hydrogen bonds to O1P of the adenine nucleotide at -7 and O2P of the same nucleotide makes a hydrogen bond to the hydroxyl of Ser B51.

The observed protein-RNA contacts arise primarily from conserved residues within the  $\beta$ -strands (Fig. 3b). Of the four residues conserved in small RNA phage coat proteins, Thr 45, Ser 47 and Pro 78 are most likely to be involved in RNA interactions or to have an important role in the assembly process. Direct contacts to the RNA originate mainly from three regions of the

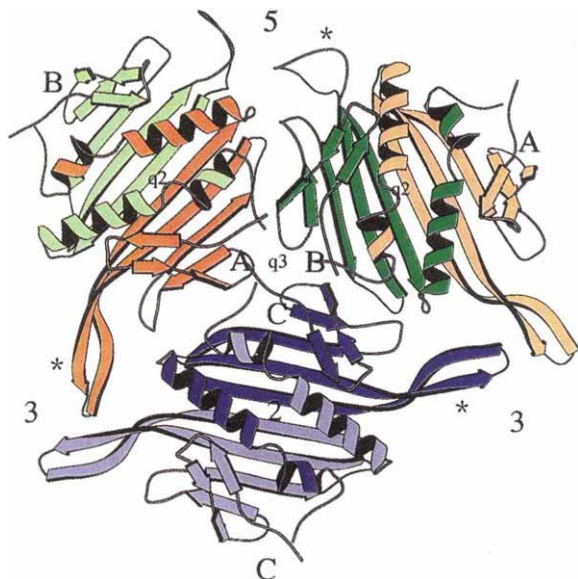


FIG. 1 Schematic drawing of the AB' and CC' dimers in one hexamer of the MS2 protein coat using the program Molscript<sup>28</sup>. Within a dimer, the monomers are arranged in such a way that the  $\beta$ -strands, five from each subunit, form a ten-stranded antiparallel  $\beta$ -sheet. The main conformational difference between the chemically identical monomers is in the long FG-loop, which makes either 5-fold or quasi-6-fold contacts. The three independent subunits are coloured orange (A), green (B) and blue (C). The symmetry-related subunits (A', B' and C') are in light colours. The view is from the outside of the particle. Icosahedral symmetry axes are indicated: 2, 2-fold; 3, 3-fold; 5, 5-fold; q2, quasi-2-fold; q3, quasi-3-fold. \* marks the FG-loop in the A, B and C subunits.

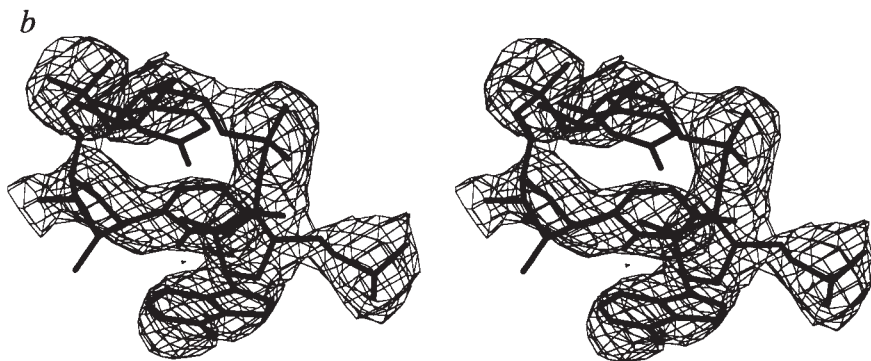
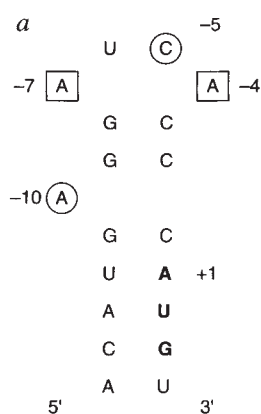
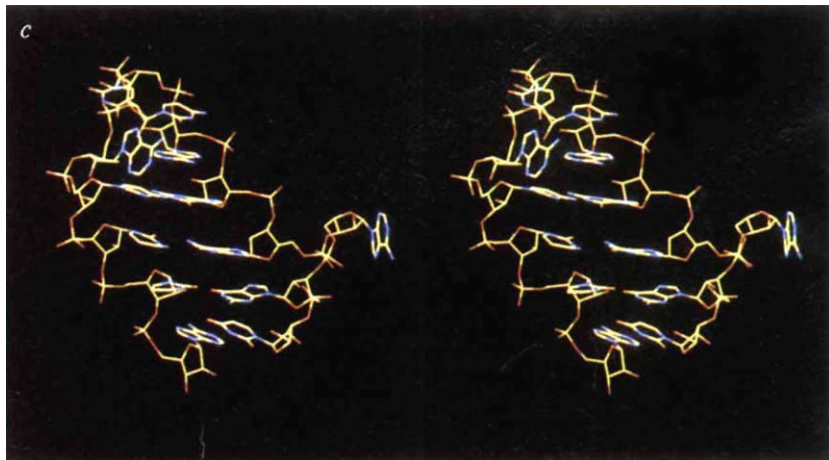


FIG. 2 a, Sequence of the 19-mer RNA fragment used for preparation of the complex. The numbering is relative to the AUG start codon (in bold) for the replicase gene at position +1. Nucleotides A -4 and A -7 (boxed) require exact identity for coat protein recognition. The -5 position and the bulged nucleotide, -10 (in circles) must be a pyrimidine and a purine, respectively, for tight binding. b, Stereo view of the fit of nucleotides A -7 to C -5 to the electron density. c, Stereo drawing of the RNA hairpin structure as bound to the AB' dimer. The RNA hairpin is composed of 13 nucleotides; a base-paired stem containing a single bulged nucleotide and a four nucleotide loop.



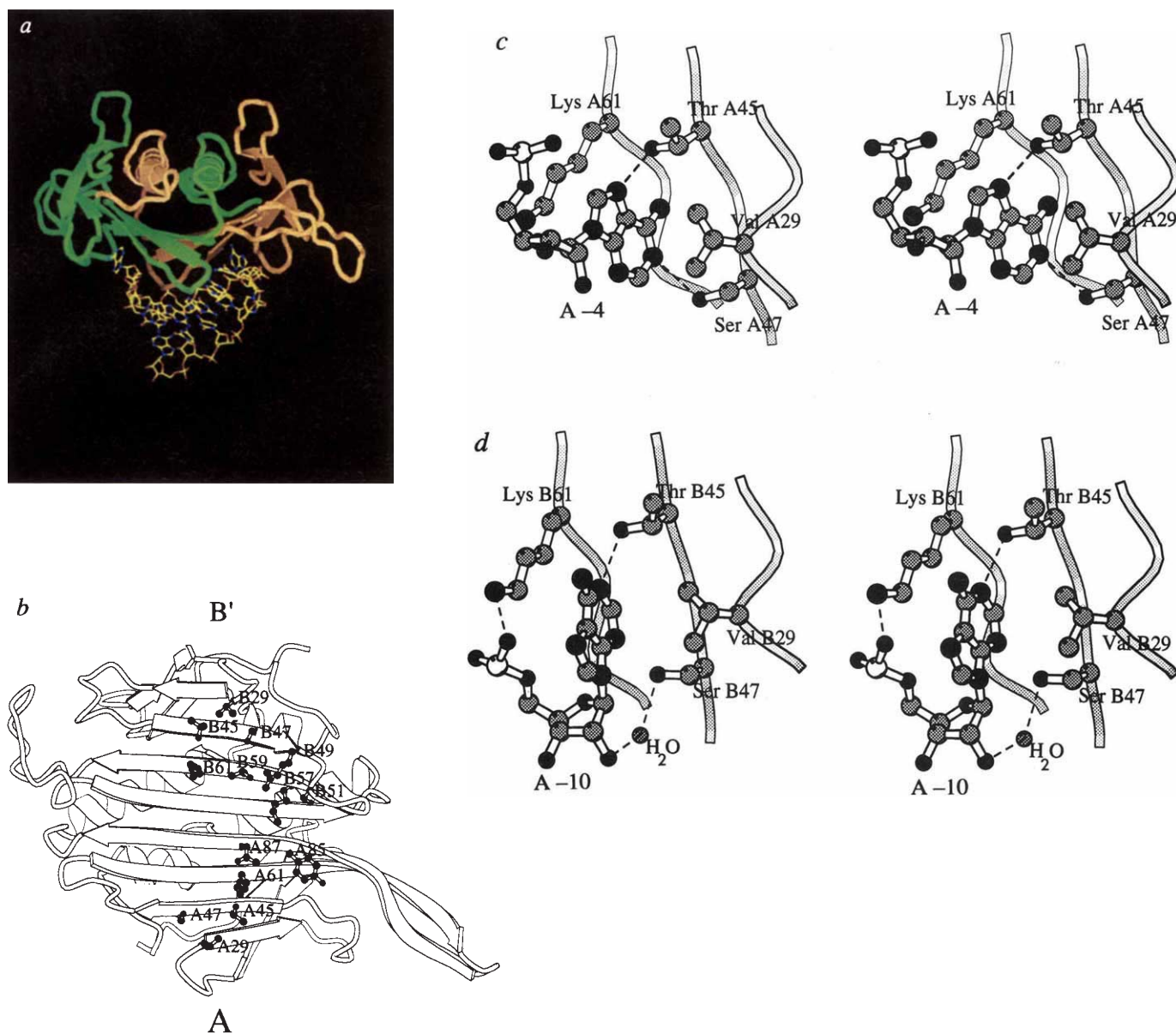
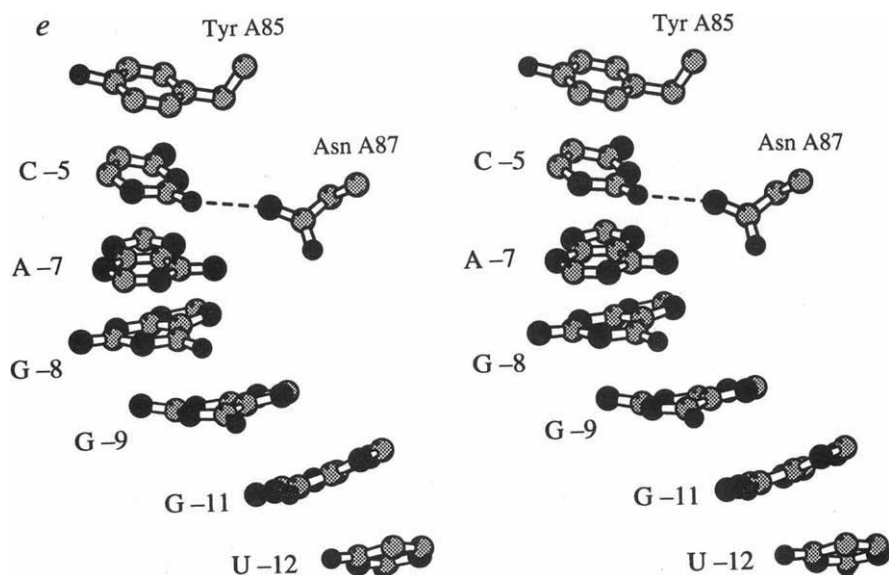


FIG. 3 *a*, Overview of the repressor-RNA complex. The RNA binds across the AB' dimer mainly at the two narrow ends of the crescent. Orange, subunit A; green, subunit B'. *b*, Molscript diagram of an AB' dimer showing amino-acid residues, which have direct protein-RNA contact. The view is from the inside of the particle. *c-e*, Molscript diagrams showing the protein interactions in the three key regions. Dashed lines define hydrogen bonds. *c*, hydroxyl groups of Thr A45 and Ser A47 make hydrogen bonds to N7 and N1 of A -4, respectively. *d*, Hydroxyl group of Thr B45 makes a hydrogen bond to N1 of A -10. The hydroxyl group of Ser B47 makes a water-mediated contact to O2\* and Lys B61 interacts with the phosphate of the adenine nucleotide at -10. *e*, Tyr A85 is stacked on the bases C -5, A -7, G -8, G -11 and U -12. The amide of Asn A87 hydrogen bonds to O2 of C -5.





dimer. The conserved residues 45 and 47 of the A and B' subunits interact with the adenines at the narrow ends of the crescent, A-4 and A-10. In both subunits, the adenines are wedged between the side chains of Val 29 and Lys 61. In the A subunit, the hydroxyl groups of Thr A45 and Ser A47 make hydrogen bonds to N7 and N1 of A-4, respectively (Fig. 3c). The interaction in the B' subunit is slightly different. The bulged nucleotide A-10 hydrogen bonds to the hydroxyl group of Thr B45, whereas the side-chain hydroxyl of Ser B47 appears to make a water mediated contact to O2\* of A-10 (Fig. 3d). Lys B61 interacts with the phosphate of the same nucleotide, which is also making a hydrogen bond to the main-chain oxygen of Thr B59.

In the MS2 operator complex, residue Tyr A85 also seems to be important for recognition. This residue is conserved within most of the *Escherichia coli* bacteriophages and Tyr 85 mutants are deficient in operator recognition<sup>9,10</sup>. Tyr A85 is stacked on the bases of C-5, A-7, G-8, G-9, G-11 and U-12 (Fig. 3e). The pyrimidine at position -5 (C-5) is bound tightly to the dimer because of hydrogen bonds O2 and N4 of C-5 and the amide of Asn A87 and O1P of the uracil nucleotide at -6, respectively.

A complete description of the correlation between the structure of the MS2 coat protein complex and the extensive biochemical data which exist for the MS2/R17 system will be published elsewhere. It is clear that RNA recognition involves direct interactions at three of the sites previously identified as essential<sup>7</sup>, namely -4, -5 and -10. Some of the residues comprising the RNA-binding site on the protein, namely Val 29, Thr 45, Lys 57, Thr 59, Lys 61, Tyr 85 and Asn 87, have been identified previously phenotypically as a result of screening mutant proteins<sup>9,10</sup>. There is no covalent adduct to the pyrimidine at -5, consistent with the properties of mutant coat proteins<sup>10,11</sup>. The requirement for A-7 may be due to its ability to participate in the tightly stacked arrangement with C-5 and G-8. The importance of Asn 87 for the recognition of a pyrimidine at position -5 has recently been demonstrated by Lim *et al.*<sup>12</sup>.

Some additional electron density is found close to Tyr B85 on the B' subunit in the dimer. It probably corresponds to the first three nucleotides in the 19-mer RNA fragment. It is possible to identify the density stacking on the tyrosine as purine (possibly A-13) but the density corresponding to the sugar-phosphate backbone in this region is very weak.

The solution conformation of the unliganded RNA operator fragment has recently been determined by NMR by Fisher and co-workers (J.F. *et al.*, manuscript in preparation). The results show that the stem is largely A helical and that the A-10 is intercalated between the neighbouring base pairs. The loop region can be described in terms of a major and minor conformation, the latter being closest to that seen in the crystal.

Little is known about the assembly process of small icosahedral viruses. For most icosahedral virus particles containing more than 60 identical subunits, a molecular switch is required to regulate the quasi-symmetry. The assembly in these viruses is regulated by an arm, which is ordered in some classes of subunits<sup>13,18</sup>. As the structure of the MS2 coat protein is completely different from these coat proteins, the assembly must be regulated in another way. The conserved Pro 78, located in the end of the FG-loop, has earlier been proposed to have a role as a switch in the assembly process<sup>10,19</sup>. In the A and C subunits, Pro 78 exists in a normal *trans* conformation but in the B subunit it has a *cis* conformation. The entire FG-loop has a completely different conformation in the B subunit, allowing the 5-fold contact to be formed. The MS2 coat protein subunits exist as dimers in solution. When the operator fragment binds to the dimer, the B conformation of the FG-loop might be stabilized and fixed by the 5'-end of the hairpin which, although partly disordered, is probably stacked to Tyr B85. The AB' dimer-RNA complex is probably part of the assembly initiation complex<sup>20</sup>. Once the assembly initiation complex is formed, more dimers could then

join the complex and their FG-loops adjusted through local protein-protein interactions.

Several RNA variants have been synthesized and similar X-ray studies are now being undertaken to give more information about the RNA recognition and virus assembly. □

Received 20 June; accepted 15 August 1994.

1. Witherell, G. W., Gott, J. M. & Uhlenbeck, O. C. *Progr. Nucleic. Acid Res. molec. Biol.* **40**, 185-220 (1991).
2. Beckett, D. & Uhlenbeck, O. C. *J. molec. Biol.* **204**, 927-938 (1988).
3. Rould, M. A., Perona, J. J. & Steitz, T. A. *Nature* **352**, 213-218 (1991).
4. Ruff, M. *et al. Science* **252**, 1682-1689 (1991).
5. Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C. & Moras, D. *Nature* **362**, 181-184 (1993).
6. Valegård, K., Liljas, L., Fridborg, K. & Unge, T. *Nature* **345**, 36-41 (1990).
7. Romaniuk, P. J., Lowary, P. T., Wu, H. N., Stormo, G. & Uhlenbeck, O. C. *Biochemistry* **26**, 1563-1568 (1987).
8. Talbot, S. J., Goodman, S., Bates, S. R. E., Fishwick, C. W. G. & Stockley, P. G. *Nucleic Acids Res.* **18**, 3521-3528 (1990).
9. Peabody, D. S. *EMBO J.* **12**, 595-600 (1993).
10. Stockley, P. G. *et al. Biochem. Soc. Trans.* **21**, 627-633 (1993).
11. Peabody, D. S. *Nucleic Acids Res.* **17**, 6017-6027 (1989).
12. Lim, F., Spingola, M. & Peabody, D. S. *J. biol. Chem.* **269**, 9006-9010 (1994).
13. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276**, 368-373 (1978).
14. Abad-Zapatero, C. *et al. Nature* **286**, 33-39 (1980).
15. Hogle, J. M., Maeda, A. & Harrison, S. C. *J. molec. Biol.* **191**, 625-638 (1986).
16. Hosur, M. V. *et al. Prot. Struct. Funct. Genet.* **2**, 167-176 (1987).
17. Liddington, R. *et al. Nature* **354**, 278-284 (1991).
18. Sorger, P. K., Stockley, P. G. & Harrison, S. C. *J. molec. Biol.* **191**, 639-658 (1986).
19. Golmohammadi, R., Valegård, K., Fridborg, K. & Liljas, L. *J. molec. Biol.* **234**, 620-639 (1993).
20. Beckett, D., Wu, H.-N. & Uhlenbeck, O. C. *J. molec. Biol.* **204**, 939-947 (1988).
21. Mastico, R. A., Talbot, S. J. & Stockley, P. G. *J. gen. Virol.* **74**, 541-542 (1993).
22. Usman, N., Ogilvie, K. K., Tiang, M.-Y. & Cedergren, R. G. *J. Am. chem. Soc.* **109**, 7845-7854 (1987).
23. Murray, J. B., Collier, A. K. & Arnold, J. R. P. *Analyt. Biochem.* **218**, 177-184 (1994).
24. Jones, T. A. in *CCP4 Study Weekend 1992; Molecular Replacement* (eds, Dodson, E. J., Gover, S. & Wolf, W.) 91-105 (Daresbury Laboratory, UK, 1992).
25. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr.* **A47**, 110-119 (1991).
26. Brünger, A. T. *XPLOR Manual* (Yale University, New Haven, CT, 1990).
27. Brünger, A. T. *Nature* **355**, 472-473 (1992).
28. Kraulis, P. J. *J. appl. Crystallogr.* **24**, 946-950 (1991).

ACKNOWLEDGEMENTS. This work was supported by the Swedish Natural Science Research Council, the UK SERC and The Wellcome Trust. We thank T. Unge for advice about the soaking procedure and J. Fisher and J. Arnold for comments on the manuscript.

## Three-dimensional structure of a peptide extending from one end of a class I MHC binding site

Edward J. Collins\*, David N. Garboczi\*†  
& Don C. Wiley†‡

\* Department of Molecular and Cellular Biology,  
† Howard Hughes Medical Institute, Harvard University,  
‡ Divinity Avenue, Cambridge, Massachusetts 02138, USA  
‡ To whom correspondence should be addressed

CLASS I major histocompatibility complex (MHC) molecules present peptides to CD8<sup>+</sup> T cells for immunological surveillance (reviewed in ref. 1). The structures of complexes of class I MHC molecules with octamer, nonamer and decamer peptides determined until now<sup>2-8</sup> show a common binding mode, with both peptide termini bound in conserved pockets at the ends of the peptide binding site. Length variations were accommodated by the peptide bulging<sup>5,6</sup> or zig-zagging<sup>7</sup> in the middle. Here we describe the structure of a decamer peptide which binds with the carboxy-terminal residue positioned outside the peptide binding site. Several protein side chains have rearranged to allow the peptide to exit. The structure suggests that even longer peptides could bind. The energetic effect of the altered mode of binding has been assessed by measuring the stability of the complex to thermal denaturation. Peptides bound in this novel manner are stable at physiological temperature, raising questions about their role in T-cell recognition and their production by proteolytic processing.

TABLE 1 Melting temperatures ( $T_m$ ) of complexes containing peptides related to MLLSVPLLLG

| Peptide     | Source       | $T_m(^{\circ}\text{C})$ | $\Delta T_m(^{\circ}\text{C})$ |
|-------------|--------------|-------------------------|--------------------------------|
| MLLSVPLLL   | Calreticulin | 65.8                    | —                              |
| MLLSVPLLLG  | Calreticulin | 51.7                    | -14.1                          |
| MLLSVPLLLGR | Experimental | 48.0                    | -17.8                          |

Each HLA-A2.1/peptide complex was purified and melting temperature assayed as described previously<sup>14</sup>. Briefly, the complexes (4–12  $\mu\text{M}$ ) were subjected to increasing temperature while the change in circular dichroic signal at 218 nm was monitored every 1.0  $^{\circ}\text{C}$  in an Aviv 62DS spectropolarimeter. The  $T_m$  was derived from the least-squares fit of the plot of molar ellipticity versus temperature using a two-state unfolding algorithm<sup>14</sup>. Data shown above are the average of two independent experiments with 2–3 melting curves per experiment. The error associated with the fitted parameter  $T_m$  is less than 0.5  $^{\circ}\text{C}$ .  $\Delta T_m$  is the change in the melting temperature from that of the nonamer MLLSVPLLL.

Following the observation that long peptides without appropriately positioned anchor residues bind class I molecules *in vivo*<sup>9,13</sup> and the suggestion that a decamer from the signal sequence of calreticulin might not bind in the carboxylate end of the binding site of HLA-A2<sup>10</sup>, we determined the structure of that decamer (MLLSVPLLLG) complexed with HLA-A2 by X-ray diffraction (Fig. 1, methods). Initial electron density maps were interpretable for the entire length of the peptide. The C-terminal glycine projects up out of the peptide binding cleft (Figs 1 and 2a). The first nine residues of the decamer bind to HLA-A2.1 almost identically to conventional nonamers; the locations of the amino terminus and the P2 anchor residue overlap those from a collection of viral peptides bound to HLA-A2 (Fig. 2b), and the P9 side chain and carbonyl oxygen overlap the P9 anchors and one oxygen of the carboxylate groups of conventionally bound nonamers (Fig. 2c). The pattern of hydrogen bonds made at the C terminus is different, however. In the conventional mode of binding (Fig. 2d), four hydrogen bonds are made to the peptide carboxylate oxygens, three from conserved MHC residues (Thr 143, Lys 146, Tyr 84) and one from a non-conserved residue (Thr 80 through  $\text{H}_2\text{O}$ ). Only one of those bonds (Thr 143) is preserved in the calreticulin decamer complex, although one new hydrogen bond is also formed between Lys 146 and the exiting P10 carboxylate group of the calreticulin decamer (Fig. 2e).

Two of the four class I residues that participate in hydrogen bonding to conventionally bound peptides have relocated in the calreticulin peptide/HLA-A2 complex. Tyr 84 has rotated away from the peptide P9 carbonyl removing the potential for a hydrogen bond (Fig. 2e). Lys 146 has raised out of the site to form a salt bridge with the P10 carboxylate. These movements create an opening through which the peptide C terminus can extend out into solvent (compare Fig. 3a and b).

We have assessed the energetic consequences of this novel binding mode by comparing the thermal stability of the calreticulin decamer complex to that of a complex with a peptide with the same first 9 amino acids (MLLSVPLLL), containing anchor residues at conventional positions (P2 Leu, P9 Leu) (Table 1). The complex with the nonamer peptide thermally denatures at 65.8  $^{\circ}\text{C}$ , a temperature comparable to the naturally occurring antigenic peptide of the influenza virus  $\text{M}_1$  protein (residues 58–66) complexed to HLA-A2<sup>14</sup>. The decamer, MLLSVPLLLG, denatures at 14  $^{\circ}\text{C}$  lower, 51.7  $^{\circ}\text{C}$ . This 14  $^{\circ}\text{C}$  difference is about half that observed (-23  $^{\circ}\text{C}$ ) when all four C-terminal hydrogen bonds (Fig. 2d) are eliminated by replacing the carboxylate group with a methyl group on the influenza virus  $\text{M}_1$  peptide<sup>14</sup>. Because the crystal structure of the calreticulin decamer suggests that longer peptides might extend even farther out of the binding site, we tested a synthetic 11-residue oligonucleotide with an arginine (a non-anchor residue for HLA-A2) added to the C

terminus of the decamer. That complex denatured at 48.0  $^{\circ}\text{C}$ , an additional loss of only 3.7  $^{\circ}\text{C}$  compared to the decamer, arguing that it also extends out of the site.

Peptides bound to class I molecules *in vivo* are derived from cellular proteins by one of two mechanisms. Most appear to arise from cytoplasmic proteolysis followed by transport into the endoplasmic reticulum (ER) (reviewed in ref. 1), but some derive from signal sequences that enter the ER lumen as part of a secreted protein<sup>9,15</sup>. The decamer (MLLSVPLLLG) studied here is the N-terminal 10 amino acids of the 17 amino-acid signal sequence of calreticulin. It is found bound to HLA-A2 even in cells lacking peptide transporter functions<sup>9</sup>. The nonamer, MLLSVPLLL, binds with substantially higher affinity than the decamer<sup>10</sup> (Table 1), yet is not detected bound to HLA-A2 *in vivo*<sup>9</sup>. These observations suggest that the nonamer is not produced *in vivo*.

It has been suggested that peptides may be trimmed<sup>16</sup>, by for example a carboxy peptidase, while bound to class I molecules. The X-ray structure (Fig. 3b) shows that the peptide carboxylate oxygens, even in this peptide which extends out of the site, are bound too closely to the MHC protein to fit simultaneously into the active site of a protease. Instead, trimming of successive

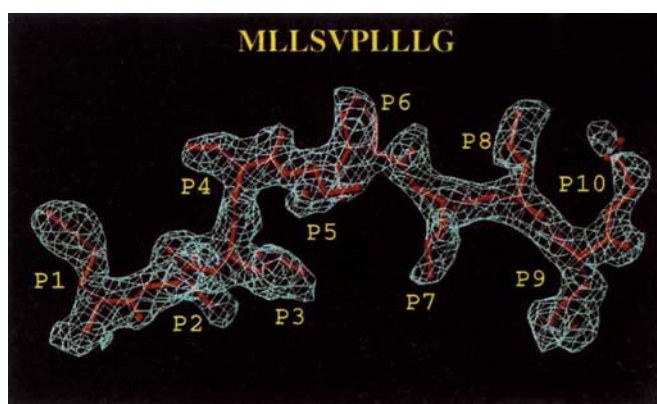


FIG. 1 'Omit'  $||2F_o| - |F_c||$  electron density for the peptide MLLSVPLLLG. The electron density is calculated from phases derived from iterative cycles of non-crystallographic real-space averaging of the two molecules in the asymmetric unit using a model without peptide. The path of the peptide is clearly visible and the glycine at P10 is located out of the peptide binding site.

**METHODS.** Soluble HLA-A2 complexed with the peptide MLLSVPLLLG was produced as described previously<sup>20</sup> and co-crystallized. X-ray data were collected to 1.9 Å resolution on supercooled crystals (-165  $^{\circ}\text{C}$ ) as 0.15° oscillation frames with a Xenotronics detector on an Elliot GX-13 rotating anode generator with a 100  $\mu\text{m}$  focal cup and focusing mirrors<sup>21</sup>. The space group was determined by examination of the differences in intensities of potential pairs of reflections across putative mirror planes within the data<sup>22</sup> and confirmed by molecular replacement and by examining the electron density maps. The space group is P1 with a unit cell of  $a = 50.41$  Å,  $b = 62.91$  Å,  $c = 74.81$  Å,  $\alpha = 81.99^{\circ}$ ,  $\beta = 76.43^{\circ}$ ,  $\gamma = 78.06^{\circ}$ . The  $R_{\text{merge}}$  for data is 5.4% from 30.0–2.07 Å, and 20.5% in the resolution shell from 2.16–2.07 Å. The molecular replacement solution was found by a rigid body fit of the reverse transcriptase peptide/HLA-A2.1 model<sup>5</sup> with X-PLOR<sup>4</sup>. The initial correlation coefficient before the rigid body fit was 64.8% from 10.0 to 2.5 Å. Cycles of real-space averaging with phase extension (G. J. Kleywegt and T. A. Jones, manuscript in preparation), manual model building using the program O<sup>23</sup>, and automated refinement using X-PLOR<sup>24</sup> were done to refine the model to 2.0 Å resolution. Although the automated refinement target was the residual based on data in the subset  $R_{\text{work}}$ , the independent parameter  $R_{\text{free}}$  (ref. 25) had to drop for the cycle to be considered successful. The final  $R_{\text{free}}$  is 29.6%, the  $R_{\text{work}}$  is 23.4% from 6.0–2.0 Å with all structure factors over 2 sigma. Each final monomer model has 251 bound waters. The geometric deviations from ideality are: r.m.s. bonds = 0.018 Å; and r.m.s. angles = 3.35 degrees.  $R_{\text{free}} = (\sum_i |F_o - F_c|) / (\sum_i F_o)$ ,  $\forall i \in \{\text{free set}\}$ ;  $R_{\text{work}} = (\sum_i |F_o - F_c|) / (\sum_i F_o)$ ,  $\forall i \in \{\text{working set}\}$ .  $R_{\text{merge}} = (\sum_{hkl} |I - \langle I \rangle|) / (\sum_{hkl} \langle I \rangle)$ ,  $\forall hkl \in \{\text{independent Miller indices}\}$ .



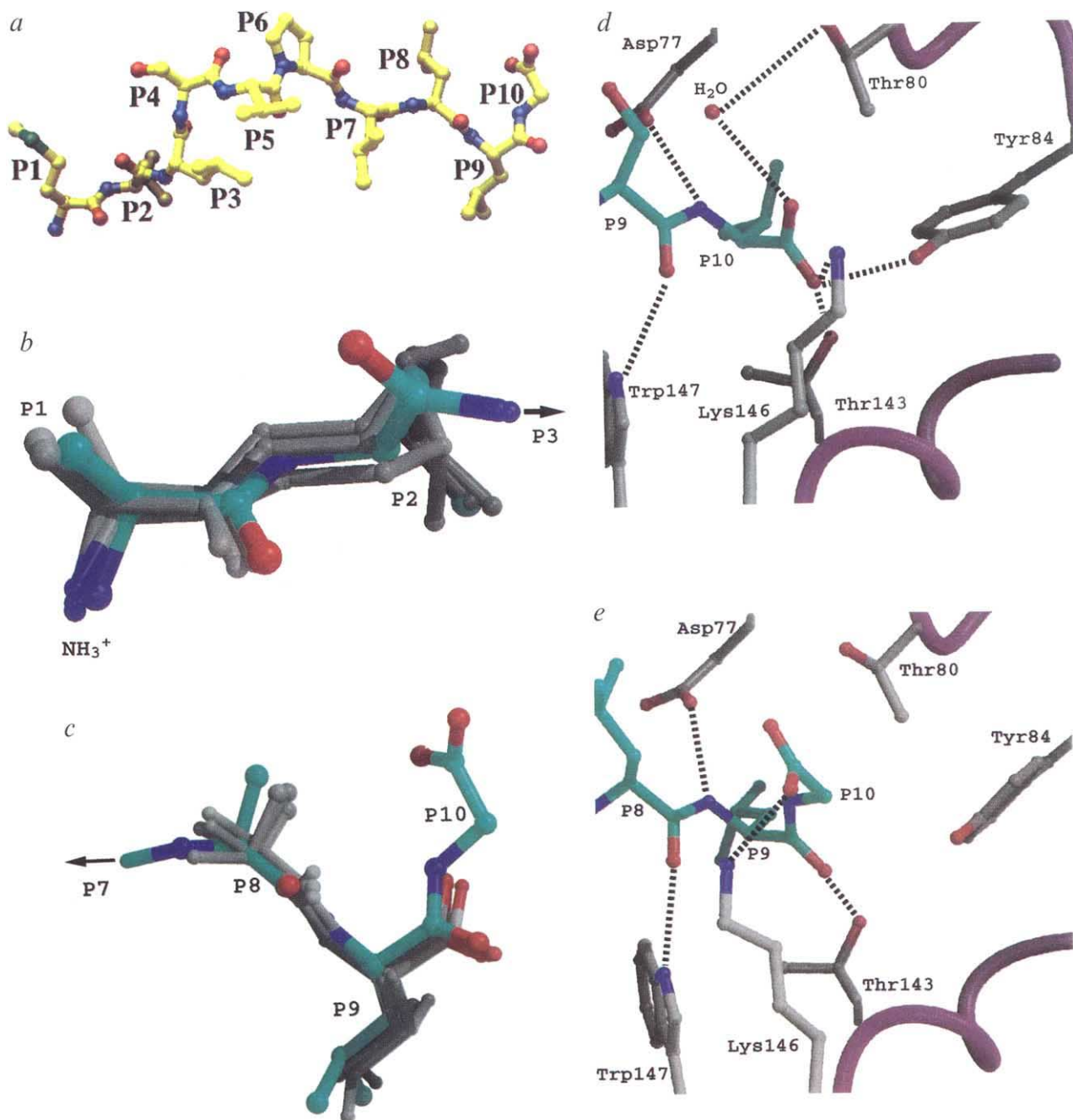


FIG. 2 The decamer, MLLSVPLLLG, binds to HLA-A2.1 in the conventional manner until residue P9. *a*, The path of the entire peptide. *b*, A superposition of the first two residues from the four nonamers previously determined<sup>4</sup> in complex with HLA-A2.1 (grey bonds) and the first two residues from MLLSVPLLLG (blue bonds). The side chains of the P1 residues have been removed for clarity (N-terminal nitrogen dark blue; P1 carbonyl oxygen red). *c*, A superposition of the last two residues of four nonamers<sup>4</sup> (grey bonds) in complex with HLA-A2.1 and the final three residues of MLLSVPLLLG complexed with HLA-A2.1 (blue bonds). The P8 side chains have been removed for clarity (carbonyl oxygens

red; amide nitrogens dark blue). *d*, An example of conventional hydrogen bonds, residues P9 and P10 of hepatitis B nucleocapsid peptide decamer (FLPSDFFPSV) bound to HLA-A2.1<sup>4</sup>. *e*, Hydrogen bonds for residues P8, P9 and P10 of calreticulin peptide (MLLSVPLLLG). *b–e*, were produced with the program HYDRASTER which was written by S. Watowich and L. Gross and based in part on programs by W. Anderson (RASTER3D) and R. Hubbard (HYDRA). The four nonamers in *b* and *c* and sources are: ILKEPVHGV (HIV-1 reverse transcriptase); LLFGYPVYV (HTLV-1 Tax); GILGFVFTL (influenza virus M peptide); TLTSNTSV (HIV-1 gp120)<sup>22</sup>.

C-terminal peptide residues might occur by a cycle of dissociation of a few C-terminal residues, proteolytic cleavage of the exposed terminal residue, and rebinding. By such a 'nibbling' mechanism a nonamer might never be produced if the decamer bound tightly enough to be transported to the cell surface with the MHC molecule.

The extraction of long peptides from class I molecules *in vivo*<sup>9–13</sup> and the thermal stability well above 37 °C measured

here, suggest that the unconventional mode of peptide binding observed crystallographically may be found on cell surfaces. It has been argued from both peptide pool sequences<sup>13</sup> and thermodynamic data<sup>14</sup>, that peptides could extend from either end of the binding site. The peptide extensions out of the binding site must be geometrically different from those found on class II MHC molecules (compare Fig. 3*a* here with Fig. 4 in ref. 18) because of differences in the shape of the binding sites<sup>17–19</sup>. It will

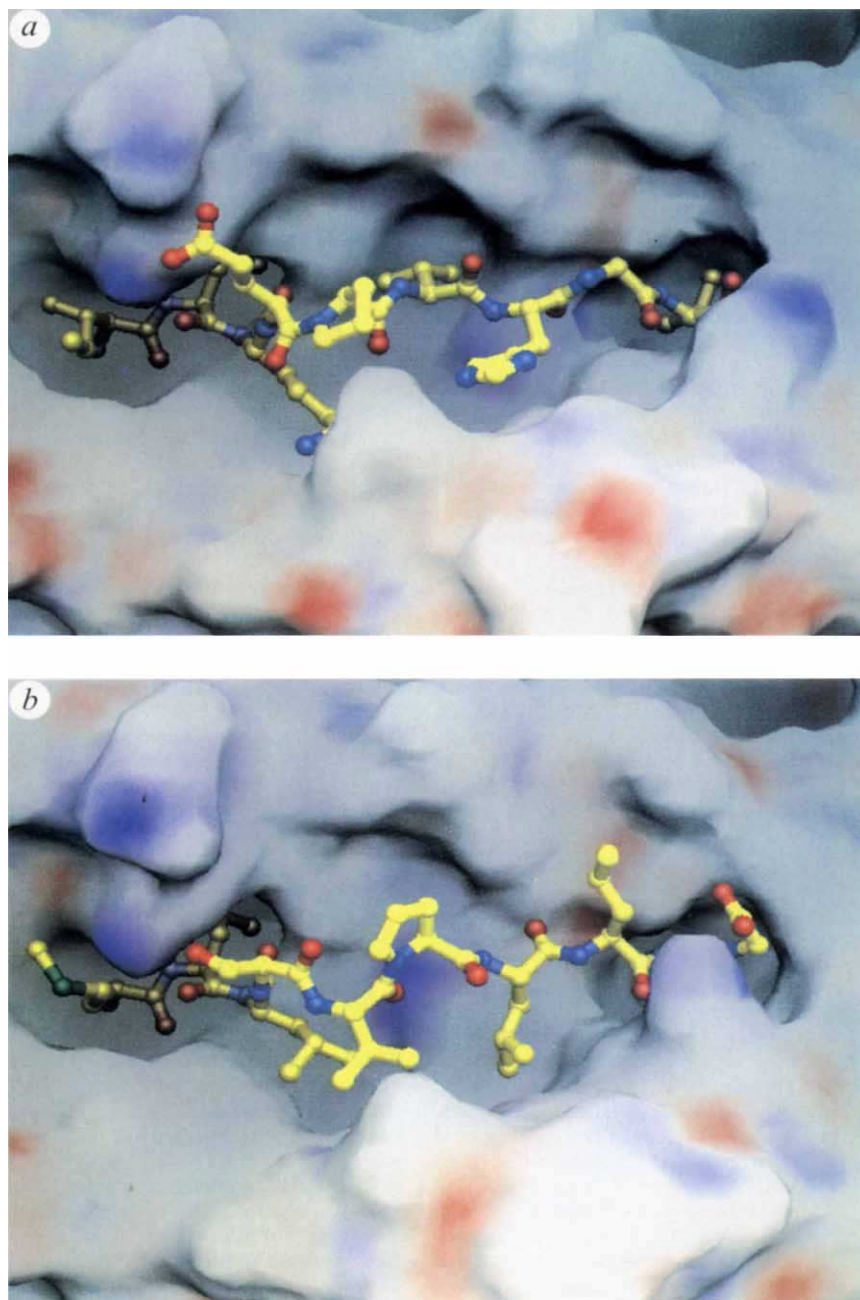


FIG. 3 The molecular surface of HLA-A2.1 complexed with a 'conventionally bound' peptide (a) and MLLSVPLLLG (b). a, HLA-A2.1 with ILKEPVHGV. The C terminus of the peptide is on the right-hand side. Lys 146 of HLA-A2.1 (visible as a positive charge, blue) forms a 'cover' under which the C terminus of the peptide is bound<sup>4</sup>. b, HLA-A2.1 with MLLSVPLLLG (calreticulin). The movement of Tyr 84 and Lys 146 (visible as a positive charge, blue) forms a groove through which the peptide C terminus exits. The electrostatic potential colouring of Figs 2a, 3a and b were produced using the program GRASP (A. Nicholls, personal communication).

be interesting to see if the novel surfaces presented by peptides extending out of the ends of class I MHC binding sites are recognized by  $\alpha\beta$  or  $\gamma\delta$  T cells. □

Received 22 June; accepted 19 August 1994.

1. Germain, R. N. & Margulies, D. H. *A. Rev. Immun.* **11**, 403–450 (1993).
2. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321–325 (1991).
3. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Cell* **70**, 1035–1048 (1992).
4. Madden, D. R., Garboczi, D. N. & Wiley, D. C. *Cell* **75**, 693–708 (1993).
5. Guo, H.-C. *et al. Nature* **360**, 364–366 (1992).
6. Zhang, W., Young, A. C. M., Imarai, M., Nathenson, S. G. & Sacchettini, J. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8403–8407 (1992).
7. Fremont, D. H., Matsumura, M., Stura, E. A., Peterson, P. A. & Wilson, I. A. *Science* **257**, 919–927 (1992).
8. Young, A. C. M., Zhang, W., Sacchettini, J. C. & Nathenson, S. G. *Cell* **76**, 39–50 (1994).
9. Henderson, R. A. *et al. Science* **255**, 1264–1266 (1992).
10. Chen, Y. *et al. J. Immun.* **152**, 2874–2881 (1994).
11. Huczko, E. L. *et al. J. Immun.* **151**, 2572–2587 (1993).
12. Olsen, A. C. *et al. Eur. J. Immun.* **24**, 385–392 (1994).

13. Urban, R. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 1534–1538 (1994).
14. Bouvier, M. & Wiley, D. C. *Science* **265**, 398–402 (1994).
15. Wei, M. & Cresswell, P. *Nature* **356**, 443–446 (1992).
16. Rotzschke, O. *et al. Nature* **348**, 252–254 (1990).
17. Brown, J. H. *et al. Nature* **364**, 33–39 (1993).
18. Stern, L. J. *et al. Nature* **368**, 215–221 (1994).
19. Stern, L. J. & Wiley, D. C. *Structure* **2**, 245–251 (1994).
20. Garboczi, D. N., Hung, D. T. & Wiley, D. C. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3429–3433 (1992).
21. Harrison, S. C. *J. Appl. Crystallogr.* **1**, 84–90 (1968).
22. Garboczi, D. N., Madden, D. R. & Wiley, D. C. *J. molec. Biol.* **239**, 581–587 (1994).
23. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr. A* **47**, 110–119 (1991).
24. Brünger, A. *X-PLOR (version 3.0) Manual* (Yale University, New Haven, Conn, 1992).
25. Brünger, A. T. *Nature* **355**, 472–475 (1992).

ACKNOWLEDGEMENTS. We thank V. H. Engelhard and Y. Chen for sharing data before publication, J. Strominger and colleagues for continued scientific collaborations on MHC molecules and L. Stern and T. Jardetzky for critical reading. E.J.C. was supported by the Irvington Institute for Medical Research and D.N.G. by a NIH National Research Service Award. D.C.W. is an Investigator of the Howard Hughes Medical Institute. The coordinates have been deposited with the Protein Data Bank, Brookhaven, New York.



# Costs of consanguinity

Analysis of the literature of the effects of consanguinity on mortality puts a figure on the price of inbreeding.

ATTITUDES to the practice of consanguinity, or marriage within the extended family, vary widely around the world. For instance, in some countries the marriage of first cousins is frowned upon, but in many cultures it is actively encouraged. In Britain, for example, only about 1 in 200 marriages are between first cousins, whereas in some parts of Africa and Asia the rate is more like 1 in 20.

In the long term, consanguineous marriages have the desirable effect of removing rare deleterious alleles from the gene pool. But the immediate consequences are that the chances of recessively inherited diseases occurring in the next generation are greatly increased. Such disorders can be debilitating if not actually fatal in themselves, for example hereditary deafness, or life-threatening conditions such as the autosomal recessive muscular dystrophies. In recent years, scientists have made many fruitful visits to places like Tunisia and Japan in search of large, inbred families affected with various rare disorders, and then applied statistical techniques such as homozygosity mapping to localize the genes in question<sup>1</sup>.

But to what extent are our negative perceptions of consanguinity coloured by these well-publicized occurrences? This question has preoccupied James Neel of the University of Michigan for almost 40 years, since, with W. J. Schull, he compiled one of the first and most comprehensive surveys of the effects of consanguinity in Japan<sup>2</sup>. In this month's *Nature Genetics*<sup>3</sup>, Neel and Alan Bittles, from the Edith Cowan University in Perth, Australia, tackle this issue by presenting a detailed analysis of 38 published studies on the effects of consanguinity on mortality.

Children of first cousins have a 1/16 chance of being homozygous at a particular locus, as a result of inheriting two identical copies of one of their grandparents' genes. By comparing the death rate of children from consanguineous marriages with that from the same population whose parents are unrelated, it is possible

to gauge just how deleterious consanguinity actually is.

For the purposes of comparison, Bittles and Neel examined death rates for children from birth (including late miscarriages and premature births after six months or more gestation) to a median age of ten years. Most of the studies they examined dealt with communities in India, Pakistan, Japan and Brazil. With just two exceptions in the 38 reports, the mortality rates among offspring of first cousins exceeded those of children of non-consanguineous unions. The combined data, say the authors, translate into a figure for the excess death rate among children of first cousins of around 4.4 per cent. For higher levels of inbreeding, this figure would in turn increase, although the degree to which such calculations are influenced by socioeconomic status (which tends to be lower among consanguineous families) is uncertain.

Neel and Bittles also suggest that this excess mortality is almost entirely due to the action of 'absolute lethals' — severe genetic disorders that prove fatal regardless of environmental factors — as opposed to 'conditional lethals', where the influence of external stresses may determine the individual's ability to survive. In a further calculation, albeit an approximate one, they derive the number of 'lethal equivalents' — the number of alleles (or combinations thereof) which produce an absolute lethal — from the mortality data. The answer is a surprisingly low 1.4.

The authors conclude by describing a curious paradox. At a rough guess, there are some 3 million (1 in 1,000) polymorphic sites in the human genome, of which 10 per cent (300,000) might lie within or close to genes and whose variation might therefore be considered potentially significant. Yet given that the estimated number of 'lethal equivalents' is orders of magnitude lower, Bittles and Neel are "forced to the conclusion that the vast majority of variation encountered in DNA is of no significance to the organism with respect to survival during a critical portion of the life span". The biggest uncertainty in all this is the degree to which inbreeding causes early fetal losses, for which there is no good measurement.

Also in this month's issue<sup>4</sup> comes characterization of the molecular basis of a new class of recessive neurological disease by Michael Litt and colleagues at the Oregon Health Sciences University. For the

first time, scientists have implicated a human potassium channel<sup>5</sup> in a hereditary disorder — episodic myokymia ataxia. Several hereditary disorders have been associated with defects in skeletal muscle, voltage-gated ion channels for sodium, calcium and chloride, including hypokalaemic periodic paralysis and paramyotonia congenita. Patients with episodic ataxia experience periodic bouts of stress-related generalized ataxia with myokymia, or rippling of the muscles.

The recent localization of the disease gene to chromosome 12p implicated a cluster of potassium channel genes. Litt and co-workers have discovered four separate missense mutations in the *KCNAl* gene in four different families, all of the mutations lying within or close to one of the six putative transmembrane regions of the channel, which is a homologue of the *Drosophila Shaker* locus. By analogy with the corresponding rat channel, the *KCNAl* protein is probably expressed in both cerebellum and peripheral nerve, in agreement with the observed phenotype.

Ion channels could well be the molecular basis of some cardiac arrhythmias, the leading cause of sudden death which claims many lives each year. Arguably the best known of this class of disorder is long QT syndrome (LQT), so called because of the lengthened interval between Q and T waves on an electrocardiogram.

It was three years ago that Mark Keating's group at the University of Utah mapped a locus for LQT to chromosome 11 (ref. 6). Now, the same group has mapped two additional LQT loci to chromosomes 7q35-36 and 3p21-24 in ten and three families, respectively<sup>7</sup>, with evidence for a fourth unmapped locus. Chloride and calcium channels are among the best contenders to be screened for these newly mapped disease genes. With the hitherto most attractive candidate for the original LQT locus, *H-ras*, now excluded by genetic linkage data<sup>8</sup>, these new localizations may be a badly needed shot in the arm for understanding LQT syndrome.

**Kevin Davies**

*Kevin Davies is Editor of Nature Genetics.*

Also in this month's *Nature Genetics*: Adeno-associated viral gene delivery to the brain; microdissecting marker chromosomes in breast cancer; localization of *IDDM3* to chromosome 15; the genetics of biotin deficiency; and the role of stop codons in alternative splicing.

1. Farrall, M. *Nature Genet.* **5**, 107-108 (1993).
2. Neel, J.V. *Physician to the Gene Pool* (Wiley, New York, 1994).
3. Bittles, A.H. & Neel, J.V. *Nature Genet.* **8**, 117-121 (1994).
4. Browne, D.L. et al. *Nature Genet.* **8**, 136-140 (1994).
5. Jan, L.Y. & Jan, Y.N. *Nature* **371**, 119-122 (1994).
6. Keating, M. et al. *Science* **252**, 704-706 (1991).
7. Jiang, C. et al. *Nature Genet.* **8**, 141-147 (1994).
8. Roy, N. et al. *Nature Genet.* **8**, 113-114 (1994).

# Small-scale magnetic isolation of mRNA and synthesis of cDNA in 96-well plates

Anne Louise Oaklander and Steven J. Ekenberg

**Multiple small biological samples can be simultaneously prepared for RT-PCR by magnetically isolating mRNA and synthesizing cDNA in 96-well plates. The resulting accuracy, high throughput and ease of use facilitate commercial and clinical applications.**

THE analysis of expression of specific mRNAs in biological samples by the reverse transcription polymerase chain reaction (RT-PCR) is increasingly common as the sequences of more genes are discovered<sup>1</sup>. Although RT-PCR is the most sensitive method of RNA analysis, it is technically difficult and has not yet achieved widespread commercial or clinical use. Amplification of a cDNA copy of an mRNA molecule is as efficient as PCR from genomic DNA, but the preceding steps in which RNA is isolated and reverse-transcribed into cDNA remain suboptimal. Biochemical methods risk loss of mRNA or contamination by DNA and are too laborious for simultaneous processing of many samples.

A magnetic method for rapidly isolating high quality polyA<sup>+</sup> mRNA and synthesizing cDNA in 96-well plates is described here. An array of ferrous pins is intermittently magnetized and moved into and out of individual wells where it manipulates mRNA attached to ferrous microparticles. Up to 96 small samples can be prepared simultaneously for PCR in a few hours. This combination of efficiency, high throughput and ease of automation makes this a useful procedure for preparing multiple cDNA samples.

## Isolation of polyA<sup>+</sup> mRNA

Compared to traditional biochemical methods, magnetic procedures for isolating mRNA are quicker and more efficient, and

they avoid organic solvents and the need for a precipitation step. Improved efficiency permits RT-PCR analysis of much smaller samples — an advantage with clinical, forensic or anthropologic specimens. This improvement is based on three strong interactions: those between adenosine (A) and thymidine (T) bases, between streptavidin and biotin, and between rare-earth magnets and paramagnetic materials (which possess no intrinsic magnetism but are attracted by a nearby magnet).

Samples are disrupted in concentrated guanidine thiocyanate (GuSCN) and  $\beta$ -mercaptoethanol to exploit their ability to dissociate nucleoproteins and inactivate ribonucleases<sup>2</sup>. Cells in suspension or growing in monolayers lyse instantly and release their mRNA into solution whereas tissues require mechanical disruption. Twenty microlitres of each homogenate (containing less than 2.5 mg of tissue or 10<sup>5</sup> cells) are transferred to a well of a 96-well plate and the remainder stored frozen for later use. High-salt buffers containing biotinylated oligo(dT) and coloured polystyrene beads are added to the lysates. The biotinylated oligo(dT) anneals strongly and specifically to repeating adenylate residues tailing the 3' end of eukaryotic mRNAs. Precipitation of proteins occurs because of the high salt concentration. After a 10-min centrifugation to pellet proteins and organelles, the polystyrene blocking beads cap the pellet (Fig. 1) and prevent it from contaminating

mRNA in the supernatant during the rest of the procedure. In traditional methods, solubilized RNA is transferred away to a fresh container for further processing. The blocking beads eliminate the need for this transfer, saving time and supplies, and eliminate an opportunity for cross-contamination or mix-up between samples, or for contamination of mRNA by the pellet.

Streptavidin-coated paramagnetic particles are then added to the supernatants and a biotin-streptavidin bond forms which holds mRNA to the particles. This complex is subsequently manipulated by the Multi-Magnet, an array of 96 ferrous pins (one for each well), which can be raised and lowered a precise distance into the wells (Fig. 1). The pins can be temporarily magnetized by sliding a rare-earth magnet over their upper ends. When the magnetized pins are lowered into the wells they bind all the mRNA complexes. The pins are withdrawn from the first plate and lowered into a fresh one containing high-stringency wash buffer. By first removing and then replacing the magnet, the mRNA complexes are released from the pins and then regathered, rinsing away all other macromolecules. The pins (and mRNA complexes) are then transferred to a third plate whose wells each contain 20  $\mu$ l of water. The absence of electrolytes causes the A:T bond to rupture and releases the purified mRNA into solution. The particles are magnetically collected and removed, leaving the

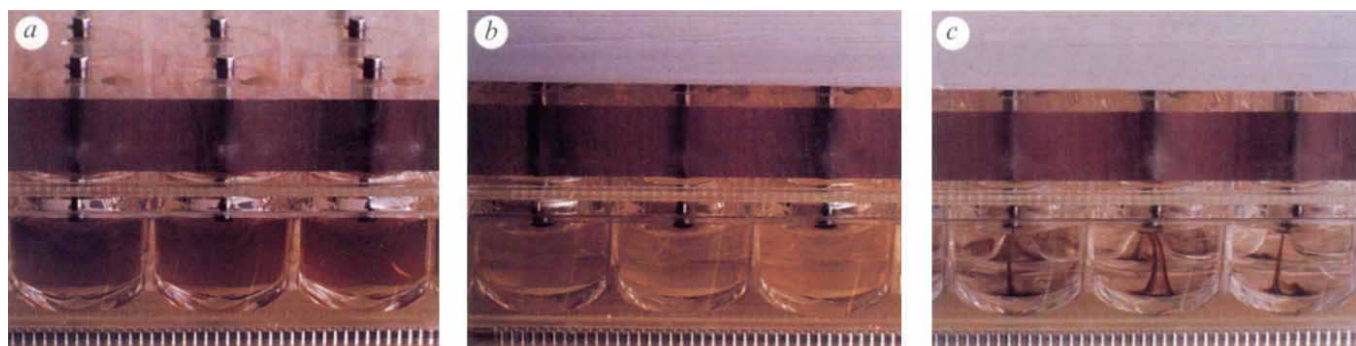


FIG. 1a, Wells after centrifugation showing capping of pellet by yellow masking beads, brown paramagnetic particles suspended in supernatants and 96-pin array in place without magnet. b, White magnet applied, mRNA-particle complex collected on tips of magnetized pins. c, Particles from the experiment in b were transferred to wash plate, released, now being recollected to magnetized pins.



purified mRNA ready for reverse transcription (RT) or other applications.

## Synthesis of cDNA

Ten microlitres of RT mix are added to each mRNA preparation. Oligo(dT) primer is used to copy all polyA<sup>+</sup> messages present. Recombinant reverse transcriptase lacking RNase-H is used to decrease degradation of the RNA template<sup>3</sup>. A novel RT buffer containing PCR-compatible dye<sup>4</sup> is used to confirm addition of RT mix to each well of mRNA, and of cDNA to each later reaction. Avoidance of pipetting errors is crucial when processing 96 samples or performing diagnostic or forensic tests. cDNA is synthesized in the same wells in which the mRNA was released (obviating another transfer). The wells are capped and incubated for 30 min to produce 30 µl of cDNA solution, enough for 6–12 fifty-microlitre PCR amplifications. The capped plates of cDNA can be frozen for future use.

## Evaluation of mRNA and cDNA

Spectrophotometric analysis of mRNA solutions demonstrated ratios of absorbance at 260 nm to 280 nm greater than or equal to 2.0, an indication that all protein had been removed. A low ratio has been shown to decrease amplification efficiency in RT-PCR<sup>5</sup>. Ratios of absorbance at 260 nm to 230 nm of 2.0 demonstrated the lack of carryover of GuSCN. Other experiments, in which radiolabelled mRNA was added to cell lysates, demonstrated an overall effi-

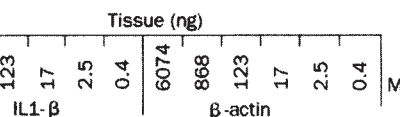


FIG. 2 Ethidium-bromide stained agarose gel of RT-PCR product prepared from sevenfold serial dilutions of mouse liver lysate. Equal volumes of the same cDNA dilutions were amplified with primers for IL1-β (left) and β-actin (right). Molecular weight markers are in the outer lanes. A signal for β-actin was obtained from as little as 2 ng of tissue, for IL1-β from as little as 123 ng. The absence of extraneous bands demonstrates high quality mRNA and cDNA, and the lack of genomic contamination.

ciency of mRNA purification greater than 90 per cent. This high efficiency allows minute samples, previously deemed inadequate, to undergo RT-PCR. Figure 2 shows that an RT-PCR signal can be obtained from as little as 2 ng of mouse liver. Dilutions of cell suspensions yielded signal from as few as two cells.

The potential contamination of mRNA preparations by genomic DNA, which can generate false positive<sup>6</sup> or negative results, has limited the use of RT-PCR. Primers that cross introns can distinguish PCR product of genomic rather than RNA origin, but are ineffective when intron position is unknown or when genes or pseudogenes lack introns. This strategy is shown in Fig. 2 for IL1-β amplification, where the presence of only one band of the expected molecular weight indicates a cDNA but not a genomic tem-

plate. Other experiments demonstrated no cross-contamination between wells, presumably because pipetting of samples was eliminated.

New technologies are needed to move PCR-based assays out of specialized laboratories into hospitals, courts, anthropology departments, and classrooms. Now that technical feasibility is established, issues of reliability, cost, ease of use and automation come to the fore. We present a novel magnetic method for preparing multiple minute samples of cDNA with improved efficiency, sensitivity and reliability. There are signs that even greater levels of miniaturization and automation are possible, such as PCR within small silicon chips<sup>7</sup>. Indeed, technological breakthroughs akin to those that fuel the microelectronics industry are possible. □

Anne Louise Oaklander is in the Department of Neurology at Johns Hopkins Medical Institutions, Baltimore, Maryland, USA and Steven J. Ekenberg is at Promega Corporation, Madison, Wisconsin, USA. We thank Paula Brisco for technical help and John W. Griffin and Steve Wesselingh for advice. This work was partly supported by the NIH and Promega. For more information, fill in reader service number 100.

1. Chelly, J. & Kahn, A. in *The Polymerase Chain Reaction* (eds Mullis, K. B., Ferre, F. & Gibbs, R. A.) 97–109 (Birkhäuser, Boston, 1994).
2. Sela, M. et al. *Biochim. biophys. Acta* **26**, 502–512 (1957).
3. Kotewicz, M. L. et al. *Nucleic Acids Res.* **16**, 265–277 (1988).
4. Hoppe, B. I. et al. *BioTechniques* **12**, 679–680 (1992).
5. Yamaguchi, M. et al. *PCR Meth. Applic.* **1**, 286–290 (1992).
6. Dilworth, D. D. & McCarrey, J. R. *PCR Meth. Applic.* **1**, 279–282 (1992).
7. Northrup, M. A. et al. in *Digest of Technical Papers: Transducers* (Proc. 7th Int. Conf. on Solid State Sensors and Actuators, Institute of Electrical and Electronic Engineers, New York, 1993).

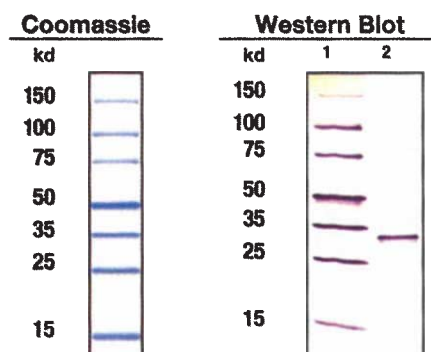
# Genetic toolkits

On offer this week — primer analysis and database searching programs, DNA purification systems, a nucleic acid quantitation reagent, PCR paraphernalia and an automated colony picking machine.

NOVAGEN has recently introduced the Perfect Protein markers for size standards on SDS-polyacrylamide gels and western blots (Reader Service No. 101). The markers, ranging in size from 15 to 150 kilodaltons, are unique in that they are all bacterially-expressed proteins. As such, Novagen says that the sizes of the proteins in the mix are

All markers carry the S•Tag peptide for one-step detection using Novagen's S-protein alkaline phosphatase conjugate. These tagged proteins can be used with any alkaline phosphatase-based detection method. Lane 1, markers detected with S-Protein and lane 2, target protein detected with antibody.

precise because they are free of oligosaccharides, which can cause anomalous migration and fuzzy bands. Each of the pro-



teins in the mix carries the S•Tag peptide which allows the marker mix to be used as size standards on western blots by including S•protein alkaline phosphatase conjugate during the second antibody incubation.

## Software

Version 3.0 of Primer Premier primer analysis software for Microsoft Windows is now available from Premier Biosoft International (Reader Service No. 102). This latest upgrade features GeneTank Premier, a new primer sequence database that permits porting of complete GenBank and EMBL sequences into the program. The new Direct Ordering Facility allows for direct ordering of selected oligonucleotides by automatic



**Premier Biosoft's primer analysis program upgrade — version 3.0.**

generation and mailing of a convenient synthesis order form via e-mail, Internet or full-colour printing. Primer Premier is fully OLE compliant, allowing import of the synthesis form into the word processor of your choice. Additional new features include restriction enzyme analysis, amino acid translation, false-priming checks and nested primer analysis. The program enables the user to weigh all search criteria, allowing optimization for specific applications. Primer selection analysis is based on the latest nearest neighbour thermodynamic theory and incorporates many advanced  $T_m$  corrections.

To keep pace with the recent advances in sequencing technology that have resulted in an explosion of nucleic acid and protein sequence data in sequence databases, IntelliGenetics and MasPar Computer announce the release of the MPSRCH 2.0 suite of **database searching programs** for the MasPar family of scalable, high-performance computers (*Reader Service No. 103*). MPSRCH contains database searching tools for discovering familial and functional relationships among primary sequence data. Authored by John Collins, joint director of the Biocomputing Research Unit at the University of Edinburgh, MPSRCH is designed specifically for database searching in



**MPSRCH: suite of database searching programs for the MasPar.**

NATURE · VOL 371 · 13 OCTOBER 1994

a high-throughput production environment. It employs the full Smith-Waterman algorithm with improved statistical analysis methods. In addition, the companies say that MPSRCH provides the broad, flexible gap handling required to retrieve sequences from the databases using a cDNA query.

The Genetics Computer Group announces the release of a new version of their **sequence analysis software**, known as the Wisconsin Package (*Reader Service No. 104*). Version 8.0 of the Wisconsin Sequence Analysis Package features the addition of a graphical user interface providing point-and-click ease of use. New programs include BLAST, which searches databases either on your local computer or at the National Center for Biotechnology Information; Prime, which helps select oligonucleotide primers; and Distances and GrowTree, which predict and display phylogenetic trees. The Wisconsin Package is a suite of programs for analysing DNA and protein sequences and searching sequence databases (GenBank, EMBL, PIR, SWISS-PROT, PROSITE, TFD and REBASE). It runs on Digital's VAX and Alpha architectures, as well as on Sun Microsystems, IBM, Silicon Graphics and Convex. Source code is included to allow users to customize applications to meet specific needs. GCG offers complete bimonthly database updates on CD-ROM; users can also keep databases current by adding incremental database updates through Internet.

### ■ Extraction

Two **DNA extraction kits** based on silica are available from Cruachem (*Reader Service No. 105*). The Isolator 1 kit contains



**Silica-based DNA extraction kits.**

sufficient reagents for up to 40 small reactions and is suitable for genomic DNA extractions from whole blood and cell cultures. The company says that the kit provides DNA with an  $A_{260/280 \text{ nm}}$  value of 1.8 and with yields typically  $\geq 10 \mu\text{g}$  of DNA from 0.5 ml blood. The kit does not require the use of Proteinase K and the whole extraction procedure can be completed in less than 2 h. For laboratories requiring larger quantities of genomic DNA, Cruachem offers the Isolator 2 kit. With this kit, sufficient reagents are provided for performing up to 50 genomic DNA extractions of 10 ml of whole blood, typically yielding 250  $\mu\text{g}$  of clean, large fragment genomic DNA from this amount of blood.

The PUREprep **TM plasmid DNA extraction kit** developed by 5 Prime to 3 Prime can now be obtained through CP Laboratories (*Reader Service No. 106*). The



**PUREprep plasmid DNA extraction kit uses no phenol.**

kit uses DNA matrix binding to yield up to 10  $\mu\text{g}$  of DNA in about 8 min, and without the need for phenol, says the company. The isolated DNA is said to be suitable for restriction enzyme digests, ligation, transfection, transformation and automated fluorescence sequencing. Each kit contains the necessary solutions, spin columns, collection tubes and DNA binding matrix.

Extract genomic DNA from cells, tissues and Gram-negative bacteria in about an hour with the four-step procedure used in the RapidGene **DNA purification kit** from Amresco (*Reader Service No. 107*). The DNA produced is said to be free of RNA contamination and has an  $A_{260/280 \text{ nm}}$  ratio of 1.7–2.0. It can be used for restriction analysis, Southern blotting, library construction and pulsed-field electrophoresis. The purity and average fragment size of the DNA produced are controlled by the user. By varying the final precipitation method, you can isolate fragments of 50–100 kilobases for restriction digestion and blotting, 100–200 kb for library construction and >200 kb for mapping and pulsed-field analysis. Typical yields are said to be 200–400  $\mu\text{g}$  from 100 mg liver tissue, 50–80  $\mu\text{g}$  from a 5 ml whole-blood sample, or 50–200  $\mu\text{g}$  from a 10-cm plate of confluent fibroblast cells.

RNeasy Total RNA kits from Qiagen are designed for rapid **RNA minipreps** of up to 100  $\mu\text{g}$  of total RNA from animal tissues,



**Total RNA minipreps without using phenol.**

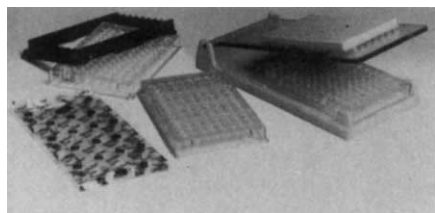


## PRODUCT REVIEW

eukaryotic cell cultures, and Gram-positive and Gram-negative bacteria (*Reader Service No. 108*). The procedure uses highly denaturing lysis conditions to generate an RNase-free environment. Total RNA is purified on an RNeasy spin column, eliminating the need for CsCl step-gradient ultracentrifugation, phenol extraction, ethanol precipitation and resuspension of the RNA pellet. Total RNA is said to be ready in 30 min for applications such as RT-PCR, northern analysis and poly A<sup>+</sup> selection.

BioPlaz is a new **reagent for the isolation of plasmid DNA** from Biogenesis (*Reader Service No. 109*). Requiring no organic extraction, columns or ultracentrifugation, BioPlaz is said to be capable of isolating plasmid DNA from 1–3 ml of culture in less than 10 min. The isolated material can be used directly for sequencing, restriction digestion, subcloning, transfection and *in vitro* transcription.

Promega's PolyAtract Series 9600 **mRNA isolation systems** are intended to allow the simultaneous isolation of polyadenylated mRNA from up to 96 tissue extracts or cell lysates in less than 2 h (*Reader Service No. 110*). The systems utilize Promega's PolyAtract technology, high-capacity MagneSphere streptavidin-coated paramagnetic particles and biotinylated oligo(dT) probe in a 96-well plate format. A



**Promega's mRNA isolation systems.**

96-pin magnetic separation array streamlines the purification process. The PolyAtract Series 9600 Systems are designed to provide an efficient means to isolate amplification-quality mRNA for developmental and tissue-specific screening and other purposes from milligram quantities of plant or animal tissues or less than 10<sup>5</sup> cells per sample. Two systems are available. One contains M-MLV reverse transcriptase and other necessary reagents for immediate conversion of the mRNA to cDNA directly in the wells of the plate. The other lacks these reagents and is intended for researchers who prefer to synthesize cDNA by other means.

A ready-to-use **reagent for the extraction of RNA-free, high molecular weight genomic DNA** up to 100 kilobases from a variety of tissue and cell culture samples is now available from Genosys (*Reader Service No. 111*). Called DNA Isolator, the reagent is provided as a chaotropic cocktail, containing no phenol. When added to the sample, DNA Isolator is said to lyse the cells



**DNA Isolator ready-to-use reagent.**

and solubilize DNA and RNA in a single step. With the addition of chloroform and mixing, a biphasic solution is created. DNA is partitioned to the aqueous phase whereas RNA and proteins remain in the organic phase. The DNA can then be isolated from the aqueous phase with isopropanol precipitation. The protocol takes less than an hour.

DNA Spinclean **columns** designed to remove low molecular weight contaminants, such as dNTPs, primers and salts, from DNA solutions are now available from Bio/Gene (*Reader Service No. 112*). Nucleic acid from 200 base pairs to 4 kilobases can be loaded onto the column in binding buffer, washed and eluted in 40 µl of tris-EDTA or water. All steps are facilitated by 2-min spins in a microcentrifuge. DNA from both TBE and TAE gels can be bound and purified on DNA Spinclean columns. Bio/Gene says that quantities ranging from >10 µg to 1 ng can be recovered from the column with yields in excess of 95 per cent.

### ■ Reagents

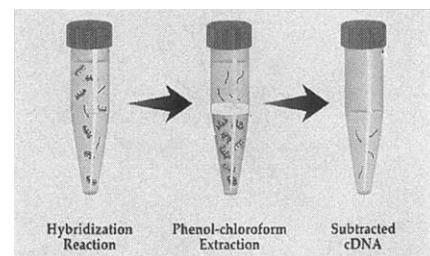
Molecular Probes has a new **reagent for quantitating double-stranded DNA** that offers an alternative to bisBenzimide-based methods (*Reader Service No. 113*). Called PicoGreen, the reagent is said to detect as little as 25 pg ml<sup>-1</sup> DNA using a standard spectrofluorometer or fluorescence microtitre plate reader equipped with fluorescein-compatible filters. In addition, the linear detection range of the assay, using a single dye concentration, extends over more than four orders of magnitude in DNA concentration, from 25 pg ml<sup>-1</sup> to 1,000 ng ml<sup>-1</sup>. The linearity of the PicoGreen DNA Quantitation assay is maintained in the presence of salts, urea, ethanol, chloroform, detergents, proteins and agarose.

Lambda terminase is an **endonuclease used for specific cleavage of DNA** in preparation for chromosomal mapping or generating restriction maps of unknown DNA cloned into cosmid vectors (*Reader Service No.*

114). It recognizes and cleaves DNA at *cos* sites, leaving, says Epicentre Technologies, unique 12-base, 5' overhangs. Cosmid restriction maps can be generated using Lambda Terminase in conjunction with a partial restriction digest, followed by hybridization with a *cos*-site-specific probe. A ladder of fragments is generated, showing the positions of restriction sites relative to each *cos* site. Epicentre's Lambda Terminase is sold in a highly purified form to provide reproducible cleavage with a minimum of nonspecific nicking activity.

The new Speed Hyb Solution from Bios Laboratories is a **ready-to-use hybridization solution** designed to reduce the time spent performing lengthy dissolving procedures and to ensure consistent results (*Reader Service No. 115*). It can be used for pre-hybridization and hybridization of membranes for Southern/northern blotting and colony and plaque screening. Two sizes are available for either 18 or 40 hybridizations.

Invitrogen's Subtractor **kit is designed for the production of subtracted cDNA** (*Reader Service No. 116*). The three-day



**Invitrogen's Subtractor kit.**

procedure uses 'subtractive hybridization' to remove common messages from a differentially expressed pool of cDNA. The subtracted cDNA produced with this kit can be used as a probe to screen a cDNA library or can be amplified and cloned to produce 'hard copies' of enriched genes. The Subtractor kit includes reagents for first-strand synthesis, biotinylation, hybridization and subtraction.

Robbins Scientific has signed a marketing agreement with Innogenetics to distribute its new line of **DNA probe-based typing kits** for the HLA/tissue typing laboratory (*Reader Service No. 117*). The InnoType DNA Typing System has been designed to eliminate the tedium and error that can be encountered with existing methods by its membrane strips containing bound DNA probes.

Novex's new 80-page **electrophoresis catalogue** for 1994/95 is available through R & D Systems (*Reader Service No. 118*). The catalogue features a variety of precast poly-



acrylamide gels for both protein and DNA separations, apparatus for running, blotting and drying gels, premixed buffers, wide-ranging molecular weight standards and staining kits.

## PCR

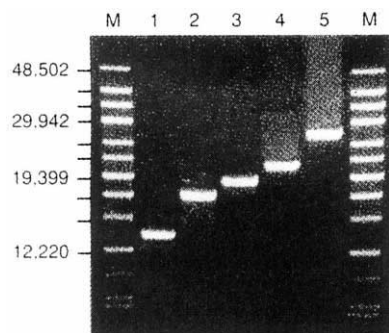
Boehringer Mannheim says that **polymerase chain reactions** can be as simple as adding primers and template to an aliquot of its new PCR Master (*Reader Service No. 119*). The premixed solution containing *Taq*



PCR Master kit — **Boehringer Mannheim.**

DNA polymerase, optimized reaction buffer and nucleotides is designed to eliminate several individual pipetting steps and thereby simplify the simultaneous preparation of multiple samples. PCR Master is also designed to guard against contamination with extraneous DNA because the premixed solution is aliquoted into ten separate vials. Each vial is said to be sufficient for ten PCR reactions. PCR Master can be stored at +4 °C, without loss of activity.

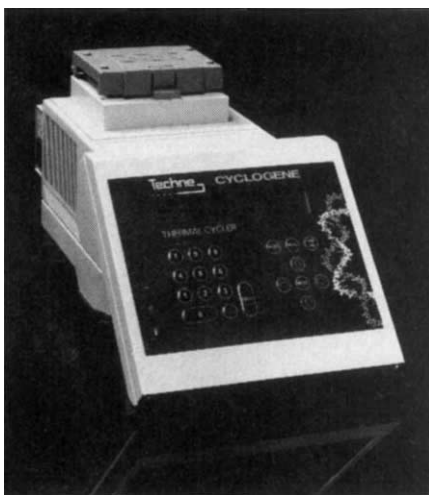
Perkin-Elmer's GeneAmp XL PCR kit is **optimized for the generation of long PCR products** — from >5 kilobases up to 40 kb (*Reader Service No. 120*). The kit uses the company's new enzyme recombinant *Thermus thermophilus* (*rTth*) DNA polymerase, XL, and a special reaction buffer. Use of the kit is said to enable the crossing of gaps in contig maps, long-range direct PCR product sequencing, crossing of intervening sequences (introns) in expressed genes and



**GeneAmp XL PCR kit. Lanes 1–3,  $\beta$ -globin cluster, human genomic DNA, 13.5, 17.7, 19.6 kb; lanes 4 and 5,  $\lambda$  DNA 20.8, 26.4 kb; M, high molecular weight marker.**

PCR-based characterization and diagnosis of medically important large gene insertions or deletions. According to Perkin-Elmer, the specially formulated *rTth* DNA polymerase, XL provides efficient extension while correcting misincorporated nucleotides (3'-to-5' exonuclease activity) that might otherwise terminate chain synthesis during subsequent annealing and extension steps.

Techné's Cyclogene **thermal cyclers** incorporate heated lids, have an automatic restart feature and can be fitted with interchangeable blocks (*Reader Service No. 121*). The heated lid enables thermal cycling to be performed without the need for mineral oil overlays and is designed to eliminate condensation and evaporation of samples. The



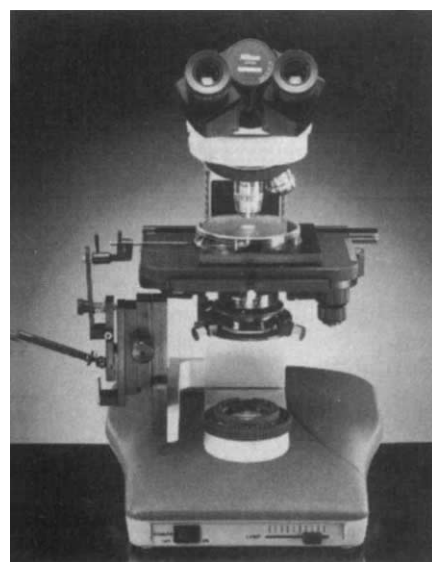
**Techné's Cyclogene thermal cycler.**

instrument also features versatile programming capabilities, a pause button for manually denaturing samples, RS232 computer output and optional sample probe. It is available in 0.5-ml, 0.2-ml, 0.25-ml, 96-well plate and flat-plate block formats.

Tropix PCR-Light is a new **system for the quantitation of PCR products** (*Reader Service No. 122*). The sensitivity of the chemiluminescent method allows detection to be performed during the exponential phase of the amplification process. The PCR-Light product is applicable in virus and RNA quantitation assays. Results are said to be obtained in about 3 h. The kit incorporates Tropix's CSPD substrate for alkaline phosphatase and Sapphire II macromolecular luminescence enhancer. The assay is performed in microplate or coated-bead assay formats.

## Instrumentation

Micro Video Instruments has a new version of its **Tetrad dissection microscope designed with yeast geneticists in mind** (*Reader Service No. 123*). Incorporating a Nikon Labophot microscope modified with a fixed stage, the instrument features a joystick micromanipulator and a specimen holder for an inverted Petri dish. Optics include planachromat  $\times 2$  and  $\times 4$  objec-



**Tetrad dissection microscope.**

tives,  $\times 15$  wide-field eyepieces and a long-working-distance condenser. A 6-V, 30-W halogen illuminator is included as standard. Additional optics and accessories are available on request. Features such as 5 mm detents installed on the y-axis of the mechanical stage are designed to enhance speed and accuracy.

Cambio is now marketing BioRobotics' BioPick, an **automated colony picker** that offers an alternative to picking colonies by hand (*Reader Service No. 124*). The process is as follows: microtitre plates have their lids removed and their wells filled using the BioFill mechanism (either 384- or 96-well plates can be used interchangeably by changing the BioFill head); they are then transported to the picking area which is a rotating turntable; each plate in turn passes under a CCD camera which scans every colony; the images are analysed on a computer, which then selects colonies of interest and transmits the positional information to the robotic



picking arm. Size, roundness and colour (blue/white) are the parameters used to determine whether a colony is suitable for picking. The colonies of interest are transferred to the appropriate microtitre wells before the next plate is passed under the camera. To reduce the likelihood of cross-contamination, the picking needle at the end of the robotic arm is sterilized between each pick. The device is designed to inoculate a



## PRODUCT REVIEW

full load of quad-density microtitre plate (9,216) wells in a day or overnight.

The UV CL-1000 crosslinkers from UVP are now available with a choice of longwave (365 nm) or midrange (302 nm) ultraviolet (UV), as well as the standard shortwave UV model (*Reader Service No. 125*). Suggested applications for longwave UV include photochemistry, solar experimentation and immunology research. For the midrange UV, applications cover phototherapy, UV curing and gradient sampling. The shortwave unit can be used in applications such as UV sterilization and UV crosslinking of DNA/RNA to membranes.

### ADVERTISEMENTS

**DNA, PEPTIDE SYNTHESIS**  
PEPTIDE/PROTEIN SEQUENCING

Personalized Service  
Guaranteed Quality

Peptides: \$26.99/residue\*  
Oligos: \$3.00/base

DNA - PEPTIDES SHIPPED  
WORLDWIDE DAILY  
PROVEN PERFORMER

**TEL: 1-519-652-3758**  
**FAX: 1-519-652-3818**

**ETROGEN CORP.**  
A Canadian Company

Reader Service No. 3

**DNA SEQUENCING**

Using an ABI 373 Automated Sequencing System  
**£60/sample**

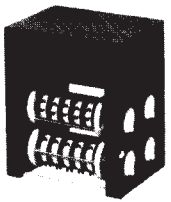
- PCR products
- M13 inserts
- plasmid inserts
- 300-400 bases of sequence
- single stranded DNA
- double stranded DNA
- solid phase sequencing

Please contact us for more information and a preparation protocol

Molecular Medicine Unit  
King's College School of Medicine & Dentistry  
Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No. 11

**BIANORM®**  
EQUILIBRIUM DIALYSIS SYSTEM



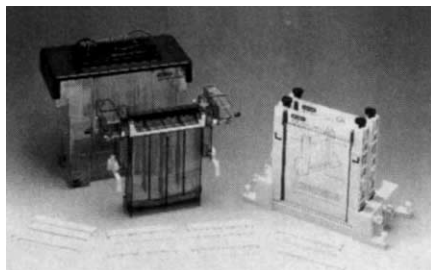
The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

**BIANORM**  
Geräte  
P.O. Box 650126  
D-81215 Munich

Reader Service No. 12

The microprocessor-controlled/UV sensor feedback system measures and controls the UV output, ensuring maximum energy efficiency.

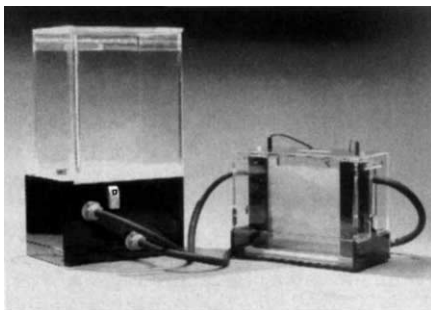
Protean II xi vertical electrophoresis cells from Bio-Rad have been designed to perform a variety of separations, including SDS-PAGE, first- and second-dimension gels for two-dimensional, native, preparative, gradient and high-resolution nucleic acid agarose electrophoresis techniques



Vertical electrophoresis cell.

(*Reader Service No. 126*). The cells allow the user to run one, two or four slab gels, or up to 16 isoelectric focusing tube gels at once. The cells are available in two gel lengths, 16 cm or 20 cm, with a selection of combs and spacers.

The new Ice-O-Therm chiller/recirculator from Owl Scientific provides researchers with a means of cooling electrophoresis gels and electroblotting transfers (*Reader Service No. 127*). Such a set-up is said to



Owl Scientific's Ice-O-Therm chiller/recirculator.

improve separation, sharpen resolution and decrease running times. The device functions by recirculating water cooled by premade ice blocks or crushed ice: no internal refrigeration is required. To operate: load ice, pour water or other cooling medium into the top-loading tank, and hook up the recirculator to a system. The Ice-O-Therm recirculator is compatible with all of the company's water-cooled vertical electrophoresis and electroblotting systems and most apparatus from other manufacturers.

### Antibodies

Berkeley Antibody Company (BAbCO) is developing novel antibodies specific for a selection of human homeobox gene

products (*Reader Service No. 128*). Currently, BAbCO has polyclonal antibodies available against four Hox proteins: HOX A10, B6, B7 and B4. These antibodies were developed using synthetic peptides and, according to the company, have been used successfully in western blots, immunoprecipitation and supershift experiments. Such research tools can be used in monitoring and studying the homeobox role in development oncology, cell differentiation and the regulation of gene expression.

Serotec has available a new monoclonal antibody to bromodeoxyuridine, which can be used in flow cytometry and immunohistology (*Reader Service No. 129*). Anti BrdU antibodies are useful reagents for assessing the proportions of proliferating cells within a population. During incubation in the presence of BrdU, actively dividing cells take up the reagent and incorporate it into their DNA, where it can be detected by using specific monoclonal antibodies. A feature of the BrdU assay is the ability to perform *in vivo* experiments. The antibody is supplied with protocols for single- and multi-colour flow cytometry and for immunohistology.

Sigma Immunochemicals has also introduced an immunohistology-grade monoclonal antibody to bromodeoxyuridine, which reacts specifically with bromodeoxyuridine (5-bromo-2-deoxyuridine, BrdU) incorporated into DNA or coupled to a protein carrier (*Reader Service No. 130*). The antibody recognizes BrdU in the nuclei of formalin-fixed, paraffin-embedded tissue sections of animals treated with *in vivo* administration of BrdU. Monoclonal anti-BrdU can be used in studies to identify DNA synthesis in cells by the detection of BrdU incorporation using immunocytochemistry and immunohistochemistry procedures. Also available is a monoclonal antibody to p53. Anti-p53 recognizes an epitope on the primate p53 nuclear protein (53 kilodaltons) that is resistant to denaturation. Sigma says that the antibody does not react with other cellular proteins. In immunoblotting, the antibody reacts with SDS-PAGE-separated whole lysates from human breast cancer cell lines that are known to express high levels of the p53 protein, but not with normal fibroblasts. Monoclonal anti-p53 should prove to be a useful tool for the detection and quantitation of p53 in malignant tissues using immunohistochemistry, immunocytochemistry, ELISAs, immunoblotting and immunoprecipitation. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details on any of the products featured, fill in the reader service card bound inside the journal. Prices quoted are sometimes nominal and apply only within the country indicated. The polymerase chain reaction (PCR) is covered by patents owned by Hoffmann-La Roche.

NATURE · VOL 371 · 13 OCTOBER 1994